ELEMENTARY

TEXT-BOOK OF PHYSICS.

BY

PROFESSOR WILLIAM A. ANTHONY,

OF CORNELL UNIVERSITY,

AND

PROFESSOR CYRUS F. BRACKETT,

OF THE COLLEGE OF NEW JERSEY.

THIRD EDITION, REVISED AND ENLARGED.

NEW YORK:

JOHN WILEY & SONS,

15 ASTOR PLACE.

1887.

H

Drummond & Neu,
Electrotypers,
1 to 7 Hague Street,
New York.

PREFACE.

THE design of the authors in the preparation of this work has been to present the fundamental principles of Physics, the experimental basis upon which they rest, and, so far as possible, the methods by which they have been established. Illustrations of these principles by detailed descriptions of special methods of experimentation and of devices necessary for their applications in the arts have been purposely omitted. The authors believe that such illustrations should be left to the lecturer, who, in the performance of his duty, will naturally be guided by considerations respecting the wants of his classes and the resources of his cabinet.

Pictorial representations of apparatus, which can seldom be employed with advantage unless accompanied with full and exact descriptions, have been discarded, and only such simple diagrams have been introduced into the text as seem suited to aid in the demonstrations. By adhering to this plan greater economy of space has been secured than would otherwise have been possible, and thus the work has been kept within reasonable limits.

A few demonstrations have been given which are not usually found in elementary text-books, except those which are much more extended in their scope than the present work. This has been done in every case in order that the argument to which the demonstration pertains may be complete and that the student may be convinced of its validity.

In the discussions the method of limits has been recognized wherever it is naturally involved; the special methods of the calculus, however, have not been employed, since, in most institutions in this country, the study of Physics is commenced before the student is sufficiently familiar with them.

The authors desire to acknowledge their obligations to Wm. F. Magie, Assistant Professor of Physics in the College of New Jersey, who has prepared a large portion of the manuscript and has aided in the final revision of all of it, as well as in reading the proof-sheets.

W. A. Anthony,
C. F. Brackett.

September, 1887.

CONTENTS.

LIGHT.

TABLES.

INTRODUCTION.

1. Divisions of Natural Science.—Everything which can affect our senses we call *matter*. Any limited portion of matter, however great or small, is called a *body*. All bodies, together with their unceasing changes, constitute *Nature*.

Natural Science makes us acquainted with the properties of bodies, and with the changes, or *phenomena*, which result from their mutual actions. It is therefore conveniently divided into two principal sections,—*Natural History* and *Natural Philosophy*.

The former describes natural objects, classifies them according to their resemblances, and, by the aid of Natural Philosophy, points out the laws of their production and development. The latter is concerned with the laws which are exhibited in the mutual action of bodies on each other.

These mutual actions are of two kinds: those which leave the essential properties of bodies unaltered, and those which effect a complete change of properties, resulting in loss of identity. Changes of the first kind are called *physical* changes; those of the second kind are called *chemical* changes. Natural Philosophy has, therefore, two subdivisions,—*Physics* and *Chemistry*.

Physics deals with all those phenomena of matter which are not directly related to chemical changes. *Astronomy* is thus a branch of Physics, yet it is usually excluded from works like the present on account of its special character.

It is not possible, however, to draw sharp lines of demarcation between the various departments of Natural Science, for the successful pursuit of knowledge in any one of them requires some acquaintance with the others.

2. Methods.—The ultimate basis of all our knowledge of nature is experience,—experience resulting from the action of bodies on our senses, and the consequent affections of our minds.

When a natural phenomenon arrests our attention, we call the result an *observation*. Simple observations of natural phenomena only in rare instances can lead to such complete knowledge as will suffice for a full understanding of them. An observation is the more complete, the more fully we apprehend the attending circumstances. We are generally not certain that all the circumstances which we note are *conditions* on which the phenomenon, in a given case, depends. In such cases we modify or suppress one of the circumstances, and observe the effect on the phenomenon. If we find a corresponding modification or failure with respect to the phenomenon, we conclude that the circumstance, so modified, is a condition. We may proceed in the same way with each of the remaining circumstances, leaving all unchanged except the single one purposely modified at each trial, and always observing the effect of the modification. We thus determine the conditions on which the phenomenon depends. In other words, we bring *experiment* to our aid in distinguishing between the real conditions on which a phenomenon depends, and the merely accidental circumstances which may attend it.

But this is not the only use of experiment. By its aid we may frequently modify some of the conditions, known to be conditions, in such ways that the phenomenon is not arrested, but so altered in the rate with which its details pass before us that they may be easily observed. Experiment also often leads to new phenomena, and to a knowledge of activities before unobserved. Indeed, by far the greater part of our knowledge of natural phenomena has been acquired by means of experiment. To be of value, experiments must be conducted with system, and so as to trace out the whole course of the phenomenon.

Having acquired our facts by observation and experiment, we seek to find out how they are related; that is, to discover the *laws* which connect them. The process of reasoning by which we discover such laws is called *induction*. As we can seldom be sure that we have apprehended all the related facts, it is clear that our inductions must generally be incomplete. Hence it follows that conclusions reached in this way are at best only probable; yet their probability becomes very great when we can discover no outstanding fact, and especially so when, regarded provisionally as true, they enable us to predict phenomena before unknown.

In conducting our experiments, and in reasoning upon them, we are often guided by suppositions suggested by previous experience. If the course of our experiment be in accordance with our supposition, there is, so far, a presumption in its favor. So, too, in reference to our reasonings: if all our facts are seen to be consistent with some supposition not unlikely in itself, we say it thereby becomes probable. The term *hypothesis* is usually employed instead of supposition.

Concerning the ultimate modes of existence or action, we know nothing whatever; hence, a law of nature cannot be demonstrated in the sense that a mathematical truth is demonstrated. Yet so great is the constancy of uniform sequence with which phenomena occur in accordance with the laws which we discover, that we have no doubt respecting their validity.

When we would refer a series of ascertained laws to some common agency, we employ the term *theory*. Thus we find in the "wave theory" of light, based on the hypothesis of a universal ether of extreme elasticity, satisfactory explanations of the laws of reflection, refraction, diffraction, polarization, etc.

3. Measurements.—All the phenomena of nature occur in *matter*, and are presented to us in *time* and *space*.

Time and space are fundamental conceptions: they do not admit of definition. Matter is equally indefinable: its distinctive characteristic is its persistence in whatever state of rest or motion it may happen to have, and the resistance which it offers to any attempt to change that state. This property is called *inertia*. It must be carefully distinguished from inactivity.

Another essential property of matter is *impenetrability*, or the property of occupying space to the exclusion of other matter.

We are almost constantly obliged, in physical science, to measure the quantities with which we deal. We measure a quantity when we compare it with some standard of the same kind. A simple number expresses the result of the comparison.

If we adopt arbitrary units of length, time, and mass (or quantity of matter), we can express the measure of all other quantities in terms of these so-called *fundamental units*. A unit of any other quantity, thus expressed, is called a *derived unit*.

It is convenient, in defining the measure of derived units, to speak of the ratio between, or the product of, two dissimilar quantities, such as space and time. This must always be understood to mean the ratio between, or the product of, the numbers expressing those quantities in the fundamental units. The result of taking such a ratio or product of two dissimilar quantities is a number expressing a third quantity in terms of a derived unit.

4. Unit of Length.—The *unit of length* usually adopted in scientific work is the *centimetre*. It is the one hundredth part of the length of a certain piece of platinum, declared to be a standard by legislative act, and preserved in the archives of France. This standard, called the *metre*, was designed to be equal in length to one ten-millionth of the earth's quadrant.

The operation of comparing a length with the standard is often difficult of direct accomplishment. This may arise from the minuteness of the object or distance to be measured, from the distant point at which the measurement is to end being inaccessible, or from the difficulty of accurately dividing our standard into very small fractional parts. In all such cases we have recourse to indirect methods, by which the difficulties are more or less completely obviated.

The *vernier* enables us to estimate small fractions of the unit of length with great convenience and accuracy. It consists of an accessory piece, fitted to slide on the principal scale of the instrument to which it is applied. A portion of the accessory piece, equal to n minus one or n plus one divisions of the principal scale, is divided into n divisions. In the former case, the divisions are numbered in the same sense as those of the principal scale; in the latter, they are numbered in the opposite sense. In either case we can measure a quantity accurately to the one nth part of one of the primary divisions of the principal scale. Fig. 1 will make the construction and use of the vernier plain.

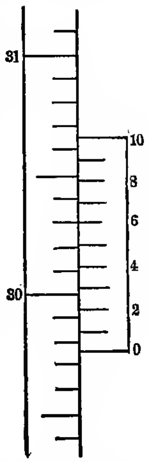

Fig. 1.

In Fig. 1, let 0, 1, 2, 3 . . . 10 be the divisions on the vernier; let 0, 1, 2, 3 . . . 10 be any set of consecutive divisions on the principal scale.

If we suppose the 0 of the vernier to be in coincidence with the limiting point of the magnitude to be measured, it is clear that, from the position shown in the figure, we have 29.7, expressing that magnitude to the nearest tenth; and since the sixth division of the vernier coincides with a whole division of the principal scale, we have $\frac{6}{10}$ of $\frac{1}{10}$, or $\frac{6}{100}$, of a principal division to be added; hence the whole value is 29.76.

The *micrometer screw* is also much employed. It consists of a carefully cut screw, accurately fitting in a nut. The head of the screw carries a graduated circle, which can turn past a fixed point. This is frequently the zero point of a scale with divisions equal in magnitude to the pitch of the screw. These divisions will then show through how many revolutions the screw is turned in any given trial; while the divisions on the graduated circle will show the fractional part of a revolution, and consequently the fractional part of the pitch that must be added. If the screw be turned through n revolutions, as shown by the scale, and through an additional fraction, as shown by the divided circle, it will pass through n times the pitch of the screw, and an additional fraction of the pitch determined by the ratio of the number of divisions read from 0 on the divided circle to the whole number into which it is divided.

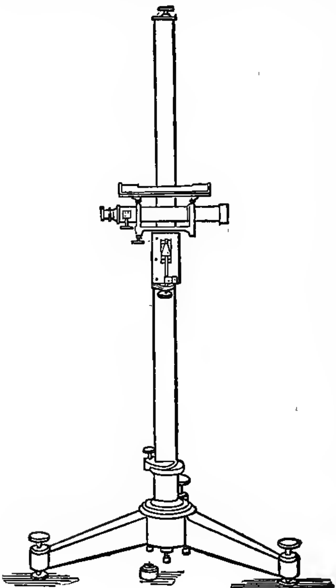

Fig. 2.

The *cathetometer* is used for measuring differences of level. A graduated scale is cut on an upright bar, which can turn about a vertical axis. Over this bar slide two accurately fitting pieces, one of which can be clamped to the bar at any point, and serve as the fixed bearing of a micrometer screw. The screw runs in a nut in the second piece, which has a vernier attached, and carries a horizontal telescope furnished with cross-hairs. The telescope having been made accurately horizontal by means of a delicate level, the cross-hairs are made to cover one of the two points, the difference of level between which is sought, and the reading upon the scale is taken; the fixed piece is then unclamped, and the telescope raised or lowered until the second point is covered by the cross-hairs, and the scale reading is again taken. The difference of scale reading is the difference of level sought.

The *dividing engine* may be used for dividing scales or for

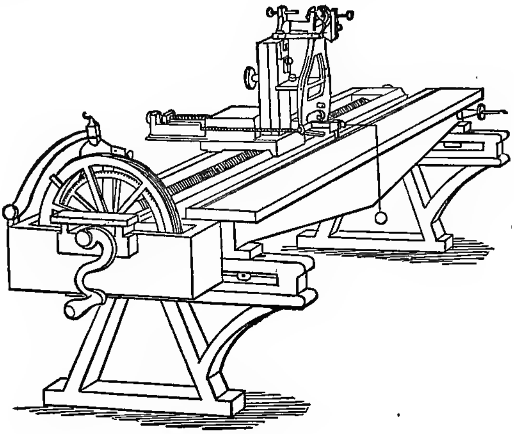

FIG. 3.

comparing lengths. In its usual form it consists essentially of a long micrometer screw, carrying a table, which slides, with a motion accurately parallel with itself, along fixed guides, resting on a firm support. To this table is fixed an apparatus for making successive cuts upon the object to be graduated.

The object to be graduated is fastened to the fixed support. The table is carried along through any required distance determined by the motion of the screw, and the cuts can be thus made at the proper intervals.

The same instrument, furnished with microscopes and accessories, may be employed for comparing lengths with a standard. It may then be called a *comparator.*

The *spherometer* is a special form of the micrometer screw. As its name implies, it is primarily used for measuring the curvature of spherical surfaces.

It consists of a screw with a large head, divided into a great number of parts, turning in a nut supported on three legs terminating in points, which form the vertices of an equilateral triangle. The axis of revolution of the screw is perpendicular to the plane of the triangle, and passes through its centre. The screw ends in a point which may be brought into the same plane with the points of the legs. This is done by placing the legs on a truly plane surface, and turning the screw till its point is just in contact with the surface. The sense of touch will enable one to decide with great nicety when the screw is turned far enough. If, now, we note the reading of the divided scale, and also that of the divided head, and then raise the screw, by turning it backward, so that the given curved surface may exactly coincide with the four points, we can compute the radius of curvature from the difference of the two readings and the known length of the side of the triangle formed by the points of the tripod.

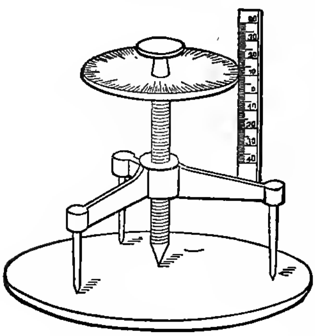

FIG. 4.

5. Unit of Time.—The *unit of time* is the *mean time second,* which is the $\frac{1}{86400}$ of a mean solar day. We employ the *clock,* regulated by the pendulum or the chronometer balance, to indicate seconds. The clock, while sufficiently accurate for ordinary use, must for exact investigations be frequently corrected by astronomical observations.

Smaller intervals of time than the second are measured by causing some vibrating body, as a tuning-fork, to trace its path along some suitable surface, on which also are recorded the beginning and end of the interval of time to be measured. The number of vibrations traced while the event is occurring determines its duration in known parts of a second.

In estimating the duration of certain phenomena giving rise to light, the revolving mirror may be employed. By its use, with proper accessories, intervals as small as forty billionths of a second have been estimated.

6. Unit of Mass.—The *unit of mass* usually adopted in scientific work is the *gram*. It is equal to the one thousandth part of a certain piece of platinum, called the *kilogram*, preserved as a standard in the archives of France. This standard was intended to be equal in mass to one cubic decimetre of water at its greatest density.

Masses are compared by means of the *balance*, the construction of which will be discussed hereafter.

7. Measurement of Angles.—Angles are usually measured by reference to a *divided circle* graduated on the system of division upon which the ordinary trigonometrical tables are based. A pointer or an arm turns about the centre of the circle, and the angle between two of its positions is measured in degrees on the arc of the circle. For greater accuracy, the readings may be made by the help of a vernier. To facilitate the measurement of an angle subtended at the centre of the circle by two distant points, a telescope with cross-hairs is mounted on the movable arm.

In theoretical discussions the unit of angle often adopted is the *radian*, that is, the angle subtended by the arc of a circle equal to its radius. In terms of this unit, a semi-circumference equals $\pi = 3.141592$. The radian, measured in degrees, is $57^\circ\ 17'\ 44.8''$.

8. Dimensions of Units.—Any derived unit may be represented by the product of certain powers of the symbols representing the fundamental units of length, mass, and time.

Any equation showing what powers of the fundamental units enter into the expression for the derived unit is called its *dimensional equation.* In a dimensional equation time is represented by T, length by L, and mass by M. To indicate the dimensions of any quantity, the symbol representing that quantity is enclosed in brackets.

For example, the unit of area varies as the square of the unit of length; hence its dimensional equation is $[\text{area}] = L^2$. In like manner, the dimensional equation for volume is $[\text{vol.}] = L^3$.

9. Systems of Units.—The system of units adopted in this book, and generally employed in scientific work, based upon the centimetre, gram, and second, as fundamental units, is called the *centimetre-gram-second system* or the C. G. S. system. A system based upon the foot, grain, and second was formerly much used in England. One based upon the millimetre, milligram, and second is still sometimes used in Germany.

MECHANICS.

CHAPTER I.

MECHANICS OF MASSES.

10. THE general subject of *motion* is usually divided, in extended treatises, into two topics,—*Kinematics* and *Dynamics*. In the first are developed, by purely mathematical methods, the laws of motion considered in the abstract, independent of any causes producing it, and of any substance in which it inheres; in the second these mathematical relations are extended and applied, by the aid of a few inductions drawn from universal experience, to the explanation of the motions of bodies, and the discussion of the interactions which are the occasion of those motions.

For convenience, the subject of Dynamics is further divided into *Statics*, which treats of forces as maintaining bodies in equilibrium and at rest, and *Kinetics*, which treats of forces as setting bodies in motion.

In this book it has been found more convenient to make no formal distinction between the mathematical relations of motion and the application of those relations to the study of forces and the motions of bodies. The subject is so extensive that only those fundamental principles and results will be presented which have direct application in subsequent parts of the work.

11. **Mass and Density.**—In many cases it is convenient to speak of the quantity of matter in a body as a whole. It is then called the *mass* of the body. In case the matter is continuously distributed throughout the body, its mass is often

represented by the help of the quantities of matter in its elementary volumes. The *density* of any substance is defined as the limit of the ratio of the quantity of matter in any volume within the substance to that volume, when the volume is diminished indefinitely. In case the distribution of matter in the body is uniform, its density may be measured by the quantity of matter in unit volume.

Since density is measured by a mass divided by a volume, its dimensions are ML^{-3}.

12. Particle.—A body constituting a part of a material system, and of dimensions such that they may be considered infinitely small in comparison with the distances separating it from all other parts of the system, is called a *particle.*

13. Motion.—The change in position of a material particle is called its *motion.* It is recognized by a change in the *configuration* of the system containing the displaced particle; that is, by a change in the relative positions of the particles making up the system. Any particle in the system may be taken as the fixed point of reference, and the motion of the others may be measured from it. Thus, for example, high-water mark on the shore may be taken as the fixed point in determining the rise and fall of the tides; or, the sun may be assumed to be at rest in computing the orbital motions of the planets. We can have no assurance that the particle which we assume as fixed is not really in motion as a part of some larger system; indeed, in almost every case we know that it is thus in motion. As it is impossible to conceive of a point in space recognizable as fixed and determined in position, our measurements of motion must always be relative.

One important limitation of this statement must be made: by proper experiments it is possible to determine the absolute angular motion of a body rotating about an axis.

14. Path.—The moving particle must always describe a continuous line or *path.* In all investigations the path may be represented by a diagram or model, or by reference to a set of assumed co-ordinates.

15. Velocity.—The rate of motion of a particle is called its *velocity*. If the particle move in a straight line, and describe equal spaces in any arbitrary equal times, its velocity is *constant*. A constant velocity is measured by the ratio of the space traversed by the particle to the time occupied in traversing that space. If s_0 and s represent the distances of the particle from a fixed point on its path at the instants t_0 and t, then its velocity is represented by

$$v = \frac{s - s_0}{t - t_0}. \tag{1}$$

If the path of the particle be curved, or if the spaces described by the particle in equal times be not equal, its velocity is *variable*. The path of a particle moving with a variable velocity may be approximately represented by a succession of very small straight lines, which, if the real path be curved, will differ in direction, along which the particle moves with constant velocities which may differ in amount. The velocity in any one of these straight lines is represented by the formula $v = \frac{s - s_0}{t - t_0}$. As the interval of time $t - t_0$ approaches zero, each of the spaces $s - s_0$ will become indefinitely small, and in the limit the imaginary path will coincide with the real path. The limit of the expression $\frac{s - s_0}{t - t_0}$ will represent the velocity of the particle along the tangent to the path at the time $t = t_0$, or, as it is called, the velocity in the path. This limit is usually expressed by $\frac{ds}{dt}$.

The practical unit of velocity is the velocity of a body moving uniformly through one centimetre in one second.

The dimensions of velocity are LT^{-1}.

16. Momentum.—The *momentum* of a body is its quantity of motion. This varies with its mass and its velocity jointly, and is measured by their product. Thus, for example, a body weighing ten grams, and having a velocity of ten centimetres, has the same momentum as a body weighing one gram, and having a velocity of one hundred centimetres. The practical unit of momentum is that of a gram of matter moving with the unit velocity. The formula is

$$mv, \tag{2}$$

where m represents mass.

The dimensions of momentum are MLT^{-1}.

17. Acceleration.—When the velocity of a particle varies, its rate of change is called the *acceleration* of the particle. Acceleration is either positive or negative, according as the velocity increases or diminishes. If the path of the particle be a straight line, and if equal changes in velocity occur in equal times, its acceleration is *constant*. It is measured by the ratio of the change in velocity to the time during which that change occurs. If v_0 and v represent the velocities of the particle at the instants t_0 and t, then its acceleration is represented by

$$f = \frac{v - v_0}{t - t_0}. \tag{3}$$

If the path of the particle be curved, or if the changes in velocity in equal times be not equal, the acceleration is *variable*. It can be easily shown, by a method similar to that used in the discussion of variable velocity, that the limit of the expression $\frac{v - v_0}{t - t_0} = \frac{dv}{dt}$ will represent the acceleration in the path at the time $t = t_0$. This acceleration is due to a change of velocity in the path. It is not in all cases the total acceleration of the particle. As will be seen in § 36, a particle moving along a curve has an acceleration which is not due to a change of velocity in the path.

The practical unit of acceleration is that of a particle, the velocity of which changes by one unit of velocity in one second.

The dimensions of acceleration are $L\ T^{-2}$.

The space $s - s_0$ traversed by a particle moving with a constant acceleration f, during a time $t - t_0$, is determined by considering that, since the acceleration is constant, the average velocity $\frac{v + v_0}{2}$ for the time $t - t_0$, multiplied by $t - t_0$, will represent the space traversed; hence

$$s - s_0 = \frac{v + v_0}{2}(t - t_0); \tag{4}$$

or, since $\frac{v}{2} = \frac{v_0 + f(t - t_0)}{2}$, we have, in another form,

$$s - s_0 = v_0(t - t_0) + \tfrac{1}{2}f(t - t_0)^2. \tag{4}$$

Multiplying equations (3) and (4), we obtain

$$v^2 = v_0^2 + 2f(s - s_0). \tag{5}$$

When the particle starts from rest, $v_0 = 0$; and if we take the starting point as the origin from which to reckon s, and the time of starting as the origin of time, then $s_0 = 0$, $t_0 = 0$, and equations (3), (4), and (5) become $v = ft$, $s = \frac{1}{2}ft^2$, and $v^2 = 2fs$.

Formula (4) may also be obtained by a geometrical construction.

At the extremities of a line AB (Fig. 5), equal in length to $t - t_0$, erect perpendiculars AC and BD, proportional to the initial and final velocities of the moving particle. For any interval of time Aa so short that the velocity during it may be considered constant, the space described is represented by the rectangle Ca, and the space described in the whole time $t - t_0$, by a point moving with a velocity increasing by successive equal increments, is represented by a series of rectangles, eb, fc, gd, etc., described on equal bases, ab, bc, cd, etc. If ab, bc . . . be diminished indefinitely, the sum of the areas of the rectangles can be made to approach as nearly as we please the area of the quadrilateral $ABCD$. This area, therefore, represents the space traversed by the point, having the initial velocity v_0, and moving with the acceleration f, through the time $t - t_0$. But $ABCD$ is equal to $AC\,(t - t_0) + (BD - AC)\,(t - t_0) \div 2$; whence

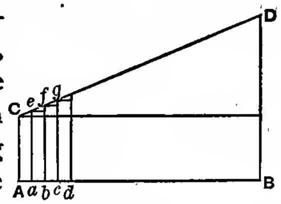

FIG. 5.

$$s - s_0 = v_0(t - t_0) + \tfrac{1}{2}f(t - t_0)^2. \qquad (4)$$

18 Composition and Resolution of Motions, Velocities, and Accelerations.—If a point a_1 move with a constant velocity relative to another point a_2, and this point a_2 move with a constant velocity relative to a third point a_3, then the motion, in any fixed time, of a_1 relative to a_3 may be readily found.

Represent the motion, in a fixed time, of a_1 relative to a_2 (Fig. 6) by the line v_2, and of a_2 relative to a_3 by the line v_3. Now, it is plain that the motions v_2 and v_3, whether acting successively or simultaneously, will bring the point a_1 to B; and also that if any portions of these motions Ab and bc, occurring in any small portion of time, be taken, they will, because the velocities of a_1 and a_2 are constant or proportional to v_2 and v_3, bring the point a_1 to some point c lying on the line joining A and

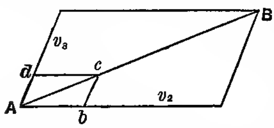

FIG. 6.

B. Therefore the diagonal *AB* of the parallelogram having the sides v_2 and v_3 fully represents the motion of a_1 relative to a_3.

The line *AB* is called the ***resultant***, of which the two lines v_2 and v_3 are the ***components***.

This proposition may now be stated generally. The resultant of any two simultaneous motions, represented by two lines drawn from the point of reference, is found by completing the parallelogram of which those lines are sides; the diagonal drawn from the point of reference represents the resultant motion.

The resultant of any number of motions may be found by obtaining the resultant of any two of the given components, by means of the parallelogram as before shown, using this resultant in combination with another component to obtain a new resultant, and proceeding in this way till all the components have been used.

The same result is reached by laying off the components as the consecutive sides of a polygon, when the line required to complete the polygon is the resultant sought.

The components of a given motion in any two given directions may be obtained by drawing lines in the two directions from one extremity of the line representing the motion, taken as origin, and constructing upon those lines the parallelogram of which the line representing the motion is the diagonal. The sides drawn from the origin represent the component motions in direction and amount.

A motion may be resolved in three directions not in the same plane by drawing from the extremity of the line representing the motion, taken as origin, lines in the three given directions, and constructing upon those lines the parallelopiped of which the line representing the motion is the diagonal. The sides of the parallelopiped drawn from the origin represent the required components.

Motions are usually resolved along three rectangular axes by means of the trigonometrical functions. Thus, if *a* be the line representing the motion, and θ, ϕ, and ψ the angles which it makes with the three axes, the components along those axes are $a \cos \theta$, $a \cos \phi$, and $a \cos \psi$.

Two motions may be compounded by first resolving them along two rectangular axes in their plane, and obtaining the resultant of the sums of their components along the axes. If a and b (Fig. 7) represent motions, $a \cos \phi$, $b \cos \theta$, $a \sin \phi$, $b \sin \theta$ are the resolved components of a and b along the axes.

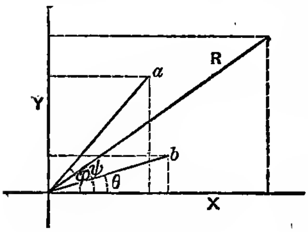

FIG. 7.

Let $a \cos \phi + b \cos \theta = X$ and $a \sin \phi + b \sin \theta = Y$; then the diagonal of the rectangle, of which X and Y are sides, is $R = (X^2 + Y^2)^{\frac{1}{2}}$; or, since the angle between the resultant and the axis of X is known by $Y = X \tan \psi$, it follows that $R = \dfrac{X}{\cos \psi}$ or $\dfrac{Y}{\sin \psi}$. It is evident that this process may be extended to any number of components in the same plane.

It is to be noted that the parallelogram law, though only proved for motions, can be shown by similar methods to be applicable to the resolution and composition of velocities and accelerations.

19. **Simple Harmonic Motion.**—If a point move in a circle with a constant velocity, the point of intersection of a diameter and a perpendicular drawn from the moving point to this diameter will have a *simple harmonic motion*. Its velocity at any instant will be the velocity in the circle resolved at that instant parallel to the diameter. The radius of the circle is the *amplitude* of the motion. The *period* is the time between any two successive recurrences of a particular condition of the moving-point. The position of a point executing a simple harmonic motion can be expressed in terms of the interval of time which has elapsed since the point last passed through the

middle of its path in the positive direction. This interval of time, when expressed as a fraction of the period, is the ***phase.***

We further define rotation in the positive direction as that rotation in the circle which is contrary to the motion of the hands of a clock, or counter-clockwise. Motion from left to right in the diameter is also considered positive. Displacement to the right of the centre is positive, and to the left negative.

If a point start from X (Fig. 8), the position of greatest positive ***elongation,*** with a simple harmonic motion, its distance s from O or its ***displacement*** at the end of the time t, during which the point in the circle has moved through the arc BX, is $OC = OB \cos \phi$. Now, OB is equal to OX, the amplitude, represented by a, and $\phi = \frac{2\pi t}{T}$, where T is the period; hence

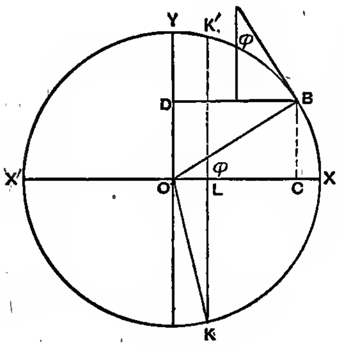

FIG. 8.

$$s = a \cos \frac{2\pi t}{T}. \qquad (6)$$

To find the velocity at the point C, we must resolve the velocity of the point moving in the circle into its components parallel to the axes. The component at the point C along OX is $V \sin \phi$; or, since $V = \frac{2\pi a}{T}$,

$$v = -\frac{2\pi a}{T} \sin \frac{2\pi t}{T}, \qquad (7)$$

remembering that motion from right to left is considered negative.

In order to find the acceleration at the point C directed towards O, we must find the rate of change of the velocity at C given by Eq. (7). Since, if the point is moving with an acceleration, the velocity increases with the time, as the time increases by a small increment Δt, the velocity also increases by the increment Δv. Eq. (7) then becomes

$$v + \Delta v = -\frac{2\pi a}{T}\sin\left(\frac{2\pi t}{T} + \frac{2\pi \Delta t}{T}\right)$$

$$= -\frac{2\pi a}{T}\left(\sin\frac{2\pi t}{T}\cos\frac{2\pi \Delta t}{T} + \cos\frac{2\pi t}{T}\sin\frac{2\pi \Delta t}{T}\right).$$

As Δt approaches zero, $\cos\frac{2\pi \Delta t}{T}$ approaches the limit unity, and $\sin\frac{2\pi \Delta t}{T}$ can be replaced by its arc $\frac{2\pi \Delta t}{T}$; making these changes, and transposing,

$$\frac{\Delta v}{\Delta t} = -\frac{4\pi^2 a}{T^2}\cos\frac{2\pi t}{T}.$$

But in the limit where these changes are admissible, $\frac{\Delta v}{\Delta t}$ becomes $\frac{dv}{dt}$; that is, the acceleration of the point.

Hence the acceleration sought is

$$f = -\frac{4\pi^2}{T^2}a\cos\frac{2\pi t}{T}. \tag{8}$$

This formula shows that the acceleration in a simple harmonic motion is proportional to the displacement. It is of the opposite sign from the displacement; that is, acceleration to the right of O is negative, and to the left of O positive.

It is often necessary to reckon time from some other position than that of greatest positive elongation. In that case the time required for the moving-point to reach its greatest positive elongation from that position, or the angle described by the corresponding point in the circumference in that time, is called the *epoch* of the new starting-point. In determining the epoch, it is necessary to consider, not only the position, but the direction of motion, of the moving-point at the instant from which time is reckoned. Thus, if L, corresponding to K in the circumference, be taken as the starting-point, the epoch is the time required to describe the path LX. But if L correspond to the point K' in the circumference, the motion in the diameter is negative, and the epoch is the time required for the moving-point to go from L through O to X' and back to X.

The epochs in the two cases, expressed in angle, are, in the first, the angle measured by the arc KX; and, in the second, the angle measured by the arc $K'X'KX$.

Choosing K in the circle, or L in the diameter, as the point from which time is to be reckoned, the angle ϕ equals angle KOB — angle KOX, or $\frac{2\pi t}{T} - \epsilon$, where t is now the time required for the moving-point to describe the arc KB, and ϵ is the epoch or the angle KOX.

The formulas then become

$$s = a\cos\left(\frac{2\pi t}{T} - \epsilon\right);$$

$$v = -\frac{2\pi}{T}a\sin\left(\frac{2\pi t}{T} - \epsilon\right);$$

$$f = -\frac{4\pi^2}{T^2}a\cos\left(\frac{2\pi t}{T} - \epsilon\right).$$

Returning to our first suppositions, letting X be the point from which epoch and time are reckoned, it is plain that, since

$$BC = a \sin \phi = a \cos\left(\phi - \frac{\pi}{2}\right) = a \cos\left(\frac{2\pi t}{T} - \frac{\pi}{2}\right),$$

the projection of B on the diameter OY also has a simple harmonic motion, differing in epoch from that in the diameter OX by $\frac{\pi}{2}$. It follows immediately that the composition of two simple harmonic motions at right angles to one another, having the same amplitude and the same period, and differing in epoch by a right angle, will produce a motion in a circle of radius a with a constant velocity. More generally, the coordinates of a point moving with two simple harmonic motions at right angles to one another are

$$x = a \cos(\phi - \epsilon) \quad \text{and} \quad y = b \cos \phi'.$$

If ϕ and ϕ' are commensurable, that is, if $\phi' = n\phi$, the curve is re-entrant. Making this supposition,

$$x = a \cos \phi \cos \epsilon + a \sin \phi \sin \epsilon, \quad \text{and} \quad y = b \cos n\phi.$$

Various values may be assigned to a, to b, and to n. Let a equal b and n equal 1; then

$$x = y \cos \epsilon + (a^2 - y^2)^{\frac{1}{2}} \sin \epsilon;$$

from which

$$x^2 - 2xy \cos \epsilon + y^2 \cos^2 \epsilon = a^2 \sin^2 \epsilon - y^2 \sin^2 \epsilon,$$

or,

$$x^2 - 2xy \cos \epsilon + y^2 = a^2 \sin^2 \epsilon.$$

This becomes, when $\epsilon = 90°$, $x^2 + y^2 = a^2$, the equation for a circle. When $\epsilon = 0°$, it becomes $x - y = 0$, the equation for a straight line through the origin, making an angle of 45° with the axis of X. With intermediate values of ϵ, it is the equation for an ellipse. If we make $n = 2$, we obtain, as special cases of the curve, a parabola and a lemniscate, according as $\epsilon = 0°$ or $90°$. If a and b are unequal, and $n = 1$, we get, in general, an ellipse.

If a line in which a point is describing a simple harmonic motion be made to move in a direction perpendicular to itself, the moving point will describe a harmonic curve, called also a sinusoid. It is a diagram of a *simple wave*. If the ordinates of the curve represent displacements transversely from a fixed line, the curve is the diagram of such waves as those of the ether which constitute light. If the ordinates of the curve represent displacement longitudinally from points of equilibrium along a fixed line, the curve may be employed to represent the waves which occur in the air when transmitting sound. The length of the wave is the distance between any two identical conditions of points on the line of progress of the wave. The amplitude of the wave is the maximum displacement from its position of equilibrium of any particle along the line of progress.

If we assume the origin of co-ordinates such that the epoch of the simple harmonic motion at the axis of ordinates is 0, the displacement from the line of progress of any point on the wave is represented by

$$s = a \cos\left(2\pi \frac{t}{T}\right).$$

The displacement due to any other wave differing from the first only in the epoch is represented by

$$s_{\prime} = a \cos\left(2\pi \frac{t}{T} - \epsilon\right).$$

We shall now show, in the simplest case, the result of compounding two wave motions.

The displacement due to both waves is the sum of the displacements due to each, hence

$$\begin{aligned} s + s_{\prime} &= a\left[\cos 2\pi \frac{t}{T} + \cos\left(2\pi \frac{t}{T} - \epsilon\right)\right] \\ &= a\left[\cos 2\pi \frac{t}{T} + \cos 2\pi \frac{t}{T} \cos \epsilon + \sin 2\pi \frac{t}{T} \sin \epsilon\right] \\ &= a\left[\cos 2\pi \frac{t}{T}(1 + \cos \epsilon) + \sin 2\pi \frac{t}{T} \sin \epsilon\right]. \end{aligned}$$

If for brevity we assume a value A and an angle ϕ such that

$$A \cos \phi = a(1 + \cos \epsilon),$$

and

$$A \sin \phi = a \sin \epsilon,$$

we may represent the last value of $s + s_{\prime}$ by

$$A \cos\left(2\pi \frac{t}{T} - \phi\right).$$

From the two equations containing A, we obtain, by adding the squares of the values of $A \sin \phi$ and $A \cos \phi$,

$$A = (2a^2 + 2a^2 \cos \epsilon)^{\frac{1}{2}};$$

and, by dividing the value of $A \sin \phi$ by that of $A \cos \phi$, we obtain

$$\phi = \tan^{-1} \frac{\sin \epsilon}{1 + \cos \epsilon}.$$

The displacement thus becomes

$$s + s_{\prime} = a(2 + 2 \cos \epsilon)^{\frac{1}{2}} \cos \left(2\pi \frac{t}{T} - \tan^{-1} \frac{\sin \epsilon}{1 + \cos \epsilon}\right). \quad (9)$$

This equation is of great value in the discussion of problems in optics.

The principle suggested by the result of the above discussion, that the resultant of the composition of two simple harmonic motions is a harmonic motion of which the elements depend on those of the components, can be easily seen to hold generally.

A very important theorem, of which this principle is the converse, was given by Fourier. It may be stated as follows: Any complex periodic function may be resolved into a number of simple harmonic functions of which the periods are commensurable with that of the original function.

As an example, any wave not simple may be decomposed into a number of simple waves the lengths of which are to each other as $\frac{1}{2}$, $\frac{1}{3}$, $\frac{1}{4}$, etc. The number of these simple waves is, in general, infinite, but in special cases determinate both as to number and to period.

20. Force.—Whenever any change occurs, or tends to occur, in the momentum of a body, we ascribe it to a cause called a *force.*

Whenever motions of matter are effected by our direct personal effort, we are conscious, through our muscular sense, of a resistance to our effort. The conception of force to which this consciousness gives rise, we transfer, by analogy, to the interaction of any bodies which is or may be accompanied by change of momentum. The question whether this analogy is or is not valid, is not involved in a purely physical discussion of the subject. A force, in the physical sense, is the assumed cause of an observed change of momentum. It is known and measured solely by the rate of that change.

If a body be moving with any acceleration whatever, the force acting on it is fully expressed by the product of the mass of the body into its acceleration.

The formula for force is, therefore,

$$\frac{mv - mv_0}{t - t_0} = mf. \tag{10}$$

The dimensions of force are MLT^{-2}.

As acceleration is always referred to some fixed direction, it follows that force is a quantity having direction.

The product of the time during which a force acts by its mean intensity is called the *impulse* of the force.

The practical unit of force is the *dyne*, which is the force that can impart to a gram of matter one unit of acceleration; that is to say, one unit of velocity in one second.

21. Field of Force.—A *field of force* is a region such that a particle constituting a part of a mutually interacting system, placed at any point in the region, will be acted on by a force, and will move, if free to do so, in the direction of the force. The particle so moving would, if it had no inertia, describe what is called a *line of force*, the tangent to which, at any point, is the direction of the force at that point. The *strength of field* at a point is measured by the force developed by unit quantity at that point, and is expressible, in terms of lines of force, by the convention that each line represents a unit of force, and that the force acting on unit quantity at any point varies as the number of lines of force which pass perpendicularly through unit area at that point. Each line, therefore, represents the direction of the force, and the number of lines passing through unit area, the strength of field. An assemblage of such lines of force considered with reference to their bounding surface is called a *tube of force*.

22. Newton's Laws of Motion.—We are now ready for the consideration of the laws of motion, first formally enunciated and successfully applied by Newton, and hence known by his name:

LEX I.—Corpus omne perseverare in statu suo quiescendi vel movendi uniformiter in directum, nisi quatenus illud a viribus impressis cogitur statum suum mutare.

LEX II.—Mutationem motus proportionalem esse vi motrici impressae & fieri secundum lineam rectam qua vis illa imprimitur.

LEX III.—Actioni contrariam semper & aequalem esse reactionem; sive corporum duorum actiones in se mutuo semper esse aequales & in partes contrarias dirigi.

The subjoined translations are given by Thomson and Tait:

LAW I.—Every body continues in its state of rest or of motion in a straight line, except in so far as it may be compelled by force to change that state.

LAW II.—Change of motion is proportional to force applied, and takes place in the direction of the straight line in which the force acts.

LAW III.—To every action there is always an equal and contrary reaction: or, the mutual actions of any two bodies are always equal and oppositely directed.

23. **Discussion of the Laws of Motion.**—(1) The first law is a statement of the important truths implied in our definition of force,—that motion, as well as rest, is a natural state of matter; that moving bodies, when entirely free to move, proceed in straight lines, and describe equal spaces in equal times; and that force is the cause of any deviation from this uniform rectilinear motion.

That a body at rest should continue indefinitely in that state seems perfectly obvious as soon as the proposition is entertained; but that a body in motion should continue to move in a straight line is not so obvious, since motions with which we are familiar are frequently arrested or altered by causes not at once apparent. This important truth, which is forced upon us by observation and experience, may, however, be presented so as to appear almost self-evident. If we conceive of a body moving in empty space, we can think of no reason why it should alter its path or its rate of motion in any way whatever.

(2) The second law presents, first, the proposition on which the measurement of force depends; and, secondly, states the identity of the direction of the change of motion with the direction of the force. Motion is here synonymous with momentum as before defined. The first proposition we have already employed in deriving the formula representing force. The second, with the further statement that more than one force can act on a body at the same time, leads directly to a most important deduction respecting the com-

bination of forces; for the parallelogram law for the resolution and composition of motions being proved, and forces being proportional to and in the same direction as the motions which they cause, it follows, if any number of forces acting simultaneously on a body be represented in direction and amount by lines, that their resultant can be found by the same parallelogram construction as that which serves to find the resultant motion. This construction is called the *parallelogram of forces.*

In case the resultant of the forces acting on a body be zero, the body is said to be in *equilibrium.*

(3) When two bodies interact so as to produce, or tend to produce, motion, their mutual action is called a *stress.* If one body be conceived as acting, and the other as being acted on, the stress, regarded as tending to produce motion in the body acted on, is a force. The third law states that all interaction of bodies is of the nature of stress, and that the two forces constituting the stress are equal and oppositely directed.

From this follows directly the deduction, that the total momentum of a system is unchanged by the interaction of its parts; that is, the momentum gained by one part is counterbalanced by the momentum lost by the others. This principle is known as the *conservation of momentum.*

24. **Collision of Bodies.**—If two bodies, m_1 and m_2, with velocities v_1 and v_2 in the same line, impinge, their velocities after contact are found, in two extreme cases, as follows:

(1) If the bodies are perfectly inelastic, there is no tendency for them to separate, their final velocities will be equal, and their momentum will be equal to the sum of their separate momenta; hence

$$m_1v_1 + m_2v_2 = (m_1 + m_2)x, \tag{11}$$

where x is the velocity after impact.

(2) If the bodies are perfectly elastic, they separate with a force equal to that by which they are compressed.

Let v represent their common velocity just at the instant when the resistance to compression balances the impulsive force. Then the change in momentum in each body up to this instant is $m_1(v_1 - v)$, or $m_2(v - v_2)$; and the further change of momentum, by reason of the elasticity of the bodies, is the same; whence the whole momentum lost by the one is $2m_1(v_1 - v)$ and that gained by the other $2m_2(v - v_2)$. If x represent the final velocity of m_1 we have the equation

$$m_1v_1 - m_1x = 2m_1v_1 - 2m_1v,$$

whence

$$x = 2v - v_1.$$

In like manner, if y represent the final velocity of m_2, we find

$$y = 2v - v_2.$$

From the formula for inelastic bodies, which is applicable at the moment when both bodies are moving with the same velocity,

$$v = \frac{m_1v_1 + m_2v_2}{m_1 + m_2};$$

whence, finally,

$$\left.\begin{aligned} x &= \frac{(m_1 - m_2)v_1 + 2m_2v_2}{m_1 + m_2}; \\ y &= \frac{(m_2 - m_1)v_2 + 2m_1v_1}{m_1 + m_2}. \end{aligned}\right\} \qquad (12)$$

25. Inertia.—The principle of equality of action and reaction holds equally well when we consider a single body as acted on by a force. The resistance to change of motion offered by the inertia of the body is equal in amount and opposite in direction to the acting force. Inertia is not of itself a force, but the property of a body, enabling it to offer a resistance to a change of motion.

26. Work and Energy.—When a force causes motion through a space, it is said to do *work.*

The measure of work is the product of the force and the space traversed by the body on which the force acts. The formula expressing work is therefore

$$mfs. \tag{13}$$

The dimensions of work are ML^2T^{-2}.

In the defined sense of the term, no work is done upon a body by a force unless it is accompanied by a change of position, and the amount of work is independent of the time taken to perform it. Both of these statements need to be made, because of a natural tendency to confound work with conscious effort, and to estimate it by the effect on our system.

A body may, in consequence of its motion or position with respect to other bodies, have a certain capacity for doing work. This capacity for doing work is its *energy.* Energy is of two kinds, usually distinguished as *potential* and *kinetic.* The former is due to the position of the body, the latter to its motion.

Since the potential energy of a body is due to the existence of a force acting upon it, it is clear that, if the body be free to move, it will be moved by the force, and its potential energy will be diminished. Hence, in any system of bodies free to move, movements will occur until the potential energy of the system becomes a minimum.

If a mass m be moving with a velocity v, its capacity for doing work may be determined from the consideration, that, if the motion be opposed by a force F equal to mf, the mass suffers a negative acceleration f, and is finally brought to rest after traversing a space s in opposition to the force. From Eq. (5) we have $s = \frac{v^2}{2f}$. Multiplying both sides of this equation by $F = mf$, we have $Fs = \frac{mv^2}{2}$. But Fs is the work done by the body against the force F, and is, therefore, the capacity which the body originally had for doing work. This capacity —that is, the kinetic energy of the body—is then represented by the expression $\frac{mv^2}{2}$.

The dimensions of energy are ML^2T^{-2}, the same as those of work. Since the square of a length cannot involve direction, it follows that energy is a quantity independent of direction.

The practical unit of work and energy is the *erg*.

It is the work done against a force of one dyne, in moving its point of application in the line of the force through a space of one centimetre;

Or, it is the energy of a body so conditioned that it can exert the force of one dyne through a space of one centimetre;

Or, it is the energy of a mass of two grams moving with unit velocity.

27. Conservation of Energy.—The difference between the kinetic energy of a body at the beginning and at the end of any given path is equal to the work done in traversing that path. For, if we consider the mass m having an acceleration f, and moving through a space s, so small that the acceleration may be assumed constant, we have, from Eq. (5),

$$v^2 = v_0^2 + 2fs,$$

where s replaces the $s - s_0$ of the equation.

Multiplying by $\frac{1}{2}m$, we have

$$\tfrac{1}{2}mv^2 = \tfrac{1}{2}mv_0^2 + mfs.$$

Since any motion whatever may be divided into portions in which the above conditions hold true, it follows that we have finally, for any motion,

$$\tfrac{1}{2}mv^2 = \tfrac{1}{2}mv_0^2 + mf_1s_1 + mf_2s_2 + \ldots = \tfrac{1}{2}mv_0^2 + \Sigma mfs.$$

Since $\frac{1}{2}mv_0^2$, or the initial kinetic energy, is a constant quantity, it follows that $\frac{1}{2}mv^2 - \Sigma mfs$, or the sum of the kinetic and potential energies, is a constant quantity for any body moving under the action of forces without collision with other bodies. In other words, a body, by losing potential energy, gains an equal amount of kinetic energy; and the kinetic energy, being used to do work against acceleration, places the body in a position where it again possesses its original amount of potential energy.

This statement holds true for any body of a system made up of bodies moving, without collision, only under their mutual interactions. It follows therefore that the total energy of such a system remains constant.

There are other forms of energy besides the potential and kinetic energies of masses. By suitable operations energy in any one form may be transformed into energy of any other form. The simplest example of such a transformation is the simultaneous production of heat and loss of mechanical energy by friction or collision.

In any *closed* system, into which no energy enters, and out of which no energy passes, the statement made above for the energy of a simple system of bodies holds true, if all forms of energy in the system are taken into account. Whatever transformations of the energy within the system occur, its total amount remains constant. This principle, called the principle of the *conservation of energy*, can be demonstrated to hold for the mechanical interaction of bodies moving without collision, and has been established by experiment for operations involving molecular and atomic interactions. It is a general principle, with which all known laws of the material universe are consistent.

The principle of the conservation of energy is so well established and so universally accepted, that, where convenient, it has been used in the demonstrations of this book as a fundamental principle.

28. Difference of Potential.—The *difference of potential* between two points in a field of force is measured by the work done by the forces of the field in moving a test unit of the quantity to the presence of which the force is due from one point to the other.

If $V_P - V_Q$ represent the difference of potential between the points P and Q, and if F represent the average force between those points and s the distance between them, then the amount of work done in moving a unit from P to Q, and hence the difference of potential between P and Q, is represented by

$$V_P - V_Q = Fs.$$

From this relation we have

$$F = \frac{V_P - V_Q}{s} = -\frac{V_Q - V_P}{s}.$$

If s become indefinitely small, in the limit F represents the force at the point P, and $-\dfrac{V_Q - V_P}{s} = -\dfrac{dV}{ds}$ becomes the rate of change of potential at that point with respect to space, taken with the opposite sign. Hence we obtain a definition of *potential*. It is a function, the rate of change of which at any point, with respect to space, taken with the opposite sign, measures the force at that point.

In the discussion which follows we deal with forces which vary directly as the product of the quantities acting, and inversely as the squares of the distances which separate them. For convenience, these acting quantities will be called masses or quantities of matter. By the substitution of proper terms the theorems to be presented will hold equally well in all cases involving forces acting according to this law.

If the field be due to the presence of a mass m, which repels the test unit, at a distance s, with a force expressed by

$$\frac{km}{s^2},$$

the difference of potential between two points, distant r and R from the mass m, is expressed by

$$V_r - V_R = km\left(\frac{1}{r} - \frac{1}{R}\right).$$

The symbol k represents the force with which two unit masses at unit distance repel one another.

To obtain this formula in the simplest case, let us suppose a mass at the point O (Fig. 9) acting upon a unit at P with a force equal to $\frac{km}{OP^2}$. If the unit be moved to Q the force at Q is $\frac{km}{OQ^2}$; and the average force acting while the unit is moving in the path PQ, provided

FIG. 9.

this path be taken small enough, is

$$\frac{km}{OP \cdot OQ}.$$

The work done in moving the unit through PQ is

$$\left(\frac{km}{OP \cdot OQ}\right)PQ = \frac{km}{OP\ OQ}(OP - OQ) = km\left(\frac{1}{OQ} - \frac{1}{OP}\right).$$

The work done in moving the unit through any other small space QR towards S is, similarly,

$$km\left(\frac{1}{OR} - \frac{1}{OS}\right).$$

The last value obtained by moving the unit from S to T is

$$km\left(\frac{1}{OT} - \frac{1}{OS}\right).$$

The sum of these values,

$$km\left(\frac{1}{OT} - \frac{1}{OP}\right),$$

gives the work done in moving the unit through the space PT.

It is evident that the amount of work done upon the unit to move it from P to T is independent of the path. For, if this were not so, it would be possible by moving from P to T on one path, and returning from T to P on another, to accumulate an indefinite amount of energy; which the principle of the conservation of energy shows to be impossible.

Since the point T can be considered as on the surface of a sphere of which O is the centre, and since the force $\frac{km}{OT^2}$ acts along TO perpendicular to that surface, and cannot, therefore, have a component tending to produce motion on that surface, there is no work done in moving the unit at T over the surface of the sphere to any other point X on it: it follows that the difference of potential between P and any other point at distance r from O is the same as that between P and T. Whence

$$V_r - V_R = km\left(\frac{1}{r} - \frac{1}{R}\right). \tag{14}$$

Such a surface as the one described, to which the lines of force are perpendicular, is called an *equipotential surface.*

If the point P be supposed to be at a distance from O so great that the force at that distance vanishes, it is then at *zero potential.* R becomes indefinitely large, and the *absolute potential* at T becomes

$$V_r = \frac{km}{r}. \tag{15}$$

This formula expresses the work necessary to move the unit against the repulsion of the mass at O up to the point T from an infinite distance. If the mass attract the unit, the work is done by the attraction upon the unit in so moving up to T, and the potential is negative.

From the definitions, it is plain that the difference of potential between P and T equals the difference of the potential energies of a unit at those points.

If the potential of any point be due to the action of more than one mass, it is found by adding the potentials due to the separate masses. If Σ be a summation sign indicating this operation, Eqs. (14) and (15) become

$$V_P - V_Q = \Sigma km\left(\frac{1}{r} - \frac{1}{R}\right), \tag{14}$$

and

$$V_r = \Sigma\frac{km}{r}.$$

29. Theorems relating to Difference of Potential.—(1) The force at any point within a spherical shell of uniform thickness and density is zero. For, if δ represent the density of the shell, and if ab_1 and cd_1 represent the volumes of the portions of the shell cut out by a cone having its apex at O (Fig. 10), then the force, if an attraction, acting towards a is $\frac{k\delta ab_1}{Oa^2}$, and towards c is $\frac{k\delta cd_1}{Oc^2}$. Hence the efficient force tending to produce motion, say towards a, is expressed by

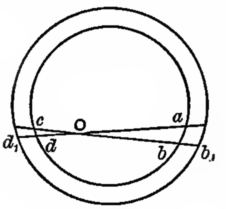

FIG. 10.

$$k\delta\left(\frac{ab_1}{Oa^2} - \frac{cd_1}{Oc^2}\right).$$

Now, if ab_1, cd_1, be taken small enough, they will be frusta of similar cones, and, as a consequence,

$$\frac{ab_1}{Oa^2} = \frac{cd_1}{Oc^2};$$

from which, since the density of the shell is uniform,

$$k\delta\left(\frac{ab_1}{Oa^2} - \frac{cd_1}{Oc^2}\right) = 0.$$

Since the whole surface of the shell may be cut by similar cones, for which similar equations will hold, the total force exerted by the shell on a unit within it becomes zero. This being so, it follows that the potential throughout the sphere is constant; for no work is required to move the unit from one point to another in the interior.

(2) The potential, and therefore the force at a point, due to the presence of a spherical shell of uniform density, depends

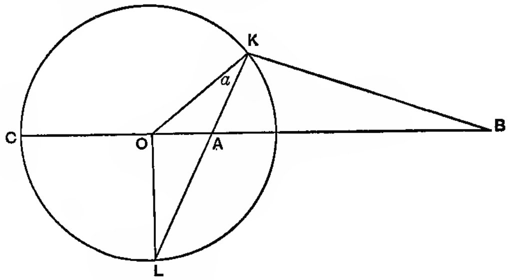

FIG. 11.

only on the mass of the shell and on the distance of the point considered from the centre of the sphere. Let *CKL* (Fig. 11) represent a central section of the shell, of which *O* is the centre. Let *d* represent the mass of a portion of the shell having unit of area. The potential at *B*, due to the element of the sphere at *K*, having an area represented by *s*, is $V_{\prime} = \frac{ds}{BK}$; and the potential due to the whole sphere is the summation of that due to all the similar elements making up the sphere. Take a point A on the line OB, such that $OA \,.\, OB = R^2$, where R is the radius of the sphere; draw AK, produce it to L, and draw OK and OL. Now, if we represent the angle OKA by α, and the solid angle subtended by the element s, as seen from A, by ω, we may express s in other terms as $\frac{AK^2 \,.\, \omega}{\cos\alpha}$; hence

$$V_{\prime} = \frac{dAK^2\omega}{BK\cos\alpha}.$$

Now, since, by construction, $OB : R = R : OA$, and the angle KOB is common to the two triangles KOA and BOK, these triangles are similar; hence

$$\frac{AK}{BK} = \frac{R}{OB}.$$

The value of the potential due to s may then be written

$$V_{\prime} = d\frac{AK}{\cos\alpha} \cdot \frac{R}{OB} \cdot \omega.$$

The value for the potential of the corresponding element at L is, similarly,

$$V_{\prime\prime} = d\frac{AL}{\cos\alpha} \cdot \frac{R}{OB} \cdot \omega.$$

Adding these values, we obtain

$$V_{\prime} + V_{\prime\prime} = d\frac{R}{OB} \cdot \omega\left(\frac{AK + AL}{\cos\alpha}\right).$$

But

$$\frac{AK + AL}{\cos\alpha} = \frac{KL}{\cos\alpha} = 2OK = 2R;$$

hence we obtain, finally,

$$V_{,} + V_{,,} = 2d\frac{R^2 \cdot \omega}{OB}.$$

Now the sphere may be divided into two portions, made up of elements similar to K and L, by a plane passing through A normal to OB. We obtain the whole potential, therefore, by summing all the potential values due to these pairs of elements; whence

$$V = 2d\frac{R^2}{OB}\Sigma\omega.$$

The sum of all the elementary solid angles on one side of the plane from the point A in it is 2π; hence, finally,

$$V = 4\pi\frac{R^2 d}{OB} = \frac{m}{OB},$$

where m is the mass of the spherical shell.

Since the force at the point B depends on the rate of change of potential at that point with respect to space, it varies inversely as the square of the distance OB. Represent OB by l. In the expression $V = \frac{m}{l}$, let l change by a small increment Δl, and denote the corresponding change in the potential by ΔV; then

$$V + \Delta V = \frac{m}{l + \Delta l},$$

$$Vl + V\Delta l + l\Delta V + \Delta V\Delta l = m.$$

If Δl become indefinitely small, in the limit the product $\Delta V \Delta l$ may be neglected. We then have

$$V\Delta l + l\Delta V = 0,$$

or

$$\frac{\Delta V}{\Delta l} = \frac{dV}{dl} = -\frac{V}{l} = -\frac{m}{l^2}.$$

This is the rate of change of potential at the point B, with respect to space, and, taken with the opposite sign, measures the force at that point. The force, therefore, at a point outside a spherical shell of uniform density varies inversely as the square of its distance from the centre of the sphere. This result enables us to deal with spheres of gravitating matter, or spherical shells, upon which is a uniform distribution of electricity, as if they were gravitating or electrified points.

(3) If in the last proposition we let $l = R$, we obtain for the value of the force just outside the shell

$$\frac{4\pi R^2 d}{R^2} = 4\pi d.$$

Since the force just inside the shell vanishes, in consequence, as we have seen, of the equal and opposite actions of the portions of the sphere ab and acb (Fig. 12), and since the total force at the point P outside the sphere is $4\pi d$, it follows that the force at P, due to ab, is $2\pi d$. If the radius be taken large enough, ab may be considered as flat, and constituting a disk: hence the force at a point near a flat disk of density d is $2\pi d$. Since the force at a point near one surface of the disk is $2\pi d$ in one direction, and near the other surface

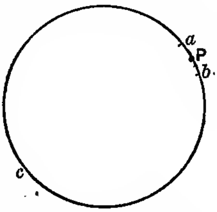

FIG. 12.

$2\pi d$ in the other direction, it is clear that, in passing through the disk, the force changes by $4\pi d$.

30. Moment of Force.—The *moment of force* about a point is defined as the product of the force and the perpendicular drawn from the point upon the line of direction of the force.

The moment of a force, with respect to a point, measures the value of the force in producing rotation about that point.

If momentum be substituted for force in the foregoing definition, we obtain the definition of *moment of momentum*.

In order to show that the moment of a force measures the value of that force in producing rotation, we will find the direction and amount of the resultant of two forces in the same plane acting on a rigid bar, but not applied at the same point.

Let BD (Fig. 13) be the bar, DF and BG the forces. Their lines of direction will, in general, meet at some point as O. Moving the forces up to O, and applying the parallelogram of forces, we obtain the resultant OJ, which cuts the bar at A. If we resolve both forces separately, parallel to OJ and BD, this resultant equals in amount the sum of those components taken parallel to OJ. Hence the components EF and CG, taken parallel to DB, annul one another's action, and, being in opposite directions, are equal. Now, by similarity of triangles,

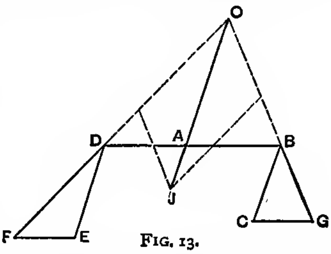

Fig. 13.

$$OA : AB = BC : CG,$$

and

$$OA : AD = DE : EF;$$

whence, since $CG = EF$, we obtain

$$AB \cdot BC = AD \cdot DE;$$

Resolving both DE and BC perpendicular to DB, we see that the moments of force about A are equal. Now, if the resultant OJ be antagonized by an equal and opposite force applied at A, there will be no motion. Hence the tendencies to rotation due to the forces are equal,—a result which is in accord with our statement that the moment of force is a measure of the value of the force in producing rotation.

The resultant of two forces may be found in general by this method. The case of most importance is the one in which the two forces are parallel. The lines DE and BC in the diagram represent such forces. It is plain, from the discussion, that these forces also will have the force represented by OJ as their resultant, applied at the point A. The resultant of two parallel forces applied at the ends of a rigid bar is then a force equal to their sum applied at a point such that the two moments of force about it are equal.

31. **Couple.**—The combination of two forces, equal and oppositely directed, acting on the ends of a rigid bar, is called a *couple.* By the preceding proposition, the resultant of these forces vanishes, and the action of a couple does not give rise to any motion of translation. The forces, however, conspire to produce rotation about the mid-point of the bar. It follows from the fact that a couple has no resultant, that it cannot be balanced by any single force.

32. **Moment of Couple.**—The *moment of couple* is the product of either of the two forces into the perpendicular distance between them. It follows from what has been already proved, that this measures the value of the couple as respects rotation.

33. **Centre of Inertia.**—If we consider any system of equal material particles, the point of which the distance from any plane whatever, is equal to the average distance of the several particles from that plane, is called the *centre of inertia.* This point is perfectly definite for any system of particles. It fol-

lows from the definition, that, if any plane pass through the centre of inertia, the sum of the distances of the particles on one side of the plane, from the plane, will be equal to the sum of the distances of the particles on the other side: hence, if the particles are all moving with a common velocity parallel to the plane, the sum of the moments of momentum on the one side is equal to the sum of the moments of momentum on the other side. And, further, if the particles all have a common acceleration, or are each acted on by equal and similarly directed forces, the sum of the moments of force on the one side is equal to the sum of the moments of force on the other side.

If we combine the forces acting on two of the particles, one on each side of the plane, we obtain a resultant equal to their sum, the distance of which from the plane is determined by the distances of the two particles from the plane. Combining this resultant with the force on another particle, we obtain a second resultant; and, by continuing this process until all the forces have been combined, we obtain a final resultant, equal to the sum of all the forces, lying in the plane, and passing through the centre of inertia. This resultant expresses, in amount, direction, and point of application, the force which, acting on a mass equal to the sum of all the particles, situated at the centre of inertia, would impart the same acceleration to it as the conjoined action of all the separate forces on the separate particles imparts to the system. When the force acting is the force of gravity, the centre of inertia is usually called the *centre of gravity*.

When the forces do not act in parallel lines, the proposition just stated does not hold true, except in special cases. Bodies in which it still holds are, for that reason, called *centrobaric* bodies.

The centre of inertia can be readily found in most of the simple geometrical figures. For the sphere, ellipsoid of revolution, or parallelopiped, it evidently coincides with the centre of figure; since a plane passing through that point in each case cuts the solid symmetrically.

34. Mechanical Powers.—The preceding definitions and propositions find their most elementary application in the so-called *mechanical powers*.

These are all designed to enable us, by the application of a certain force at one point, to obtain at another point a force, in general not equal to the one applied. Six mechanical powers are usually enumerated,—the lever, pulley, wheel and axle, inclined plane, wedge, and screw.

(1) *The Lever* is any rigid bar, of which the weight may be neglected, resting on a fixed point called a *fulcrum*. From the proposition in § 30, it may be seen, that, if forces be applied to the ends of the lever, there will be equilibrium when the resultant passes through the fulcrum. In that case the moments of force about the fulcrum are equal; whence, if the forces act in parallel lines, it follows that the force at one end is to the force at the other end in the inverse ratio of the lengths of their respective lever-arms. If l and $l_{\prime}$ represent the lengths of the arms of the lever, and P and $P_{\prime}$ the forces applied to their respective extremities, then $Pl = P_{\prime}l_{\prime}$.

The principle of the equality of action and re-action enables us to substitute for the fulcrum a force equal to the resultant of the two forces. We have then a combination of forces as represented in the diagram (Fig. 14). Plainly any one of these forces may be considered as taking the place of the fulcrum, and either of the others the *power* or the *weight*.

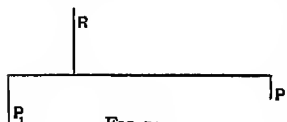

FIG. 14.

The lever is said to be of the first kind if R is fulcrum and P power, of the second kind if $P_{\prime}$ is fulcrum and P power, of the third kind if P is fulcrum and R power.

(2) *The Pulley* is a frictionless wheel, in the groove of which runs a perfectly flexible, inextensible cord.

If the wheel be on a fixed axis, the pulley merely changes the direction of the force applied at one end of the cord. If the wheel be movable and one end of the cord fixed, and a force be applied to the other end parallel to the direction of the first part of the cord, the force acting on the pulley is double the force applied: for the stress on the cord gives rise to a force in each branch of it equal to the applied force; each of these forces acts on the wheel, and, since the radii of the wheel are equal, the resultant of these two forces is a force equal to their sum applied at the centre of the wheel. From these facts the relation of the applied force to the force obtained in any combination of pulleys is evident.

(3) *The Inclined Plane* is any frictionless surface, making an angle with the line of direction of the force applied at a point upon it. Resolving the force P (Fig. 15), making an angle ϕ with the normal to the plane, into its components $P \cos \phi$ and $P \sin \phi$ perpendicular to and parallel with the plane, $P \sin \phi$ is alone effective to produce motion. Consequently, a force $P \sin \phi$ acting parallel to the surface will balance a force P, making an angle ϕ with the normal to the surface. If the plane be taken as the hypothenuse of a right-angled triangle ABC, of which the base AB is perpendicular to the line of direction of the force, then, by similarity of triangles, the angle BAC equals ϕ: whence the force obtained parallel to AC is equal to the force applied multiplied by the sine of the angle of inclination of the plane. If the components of the force applied be taken, the one, as before, perpendicular to the plane AC, and the other parallel to the base AB, the force obtained parallel to AB is equal to the force applied multiplied by the tangent of the angle of inclination of the plane.

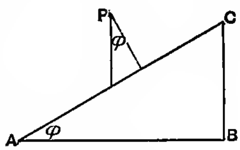

FIG. 15.

(4) *The Wheel and Axle* is essentially a continuously acting lever.

(5) *The Wedge* is made up of two similar inclined planes set together, base to base.

(6) *The Screw* is a combination of the lever and the inclined plane.

The special formulas expressing the relations of the force applied to the force obtained by the use of these combinations, are deduced from those for the more elementary mechanical powers.

It may be seen, in general, in the use of the mechanical powers, that the force applied is not equal to the force obtained. A little consideration will show, however, that the energy expended is always equal to the work done.

Any arrangement of the mechanical powers, designed to do work, is called a *machine*. The more nearly the value of the work done approaches that of the energy expended, the more closely the machine approaches perfection. The elasticity of the materials we are compelled to employ, friction, and other causes which modify the conditions required by theory, make the attainment of such perfection impossible.

The ratio of the useful work done to the energy expended is called the *efficiency* of the machine. Since in every actual machine there is a loss of energy in the transmission, the efficiency is always a proper fraction.

35. **Angular Velocity.**—The angle contained by the line passing through two points, one of which is in motion, and any assumed line passing through the fixed point, will, in general, vary. The rate of its change is called the *angular velocity* of the moving-point. If ϕ and ϕ_0 represent the angles made by the moving line with the fixed line at the instants t and t_0, then the angular velocity, if constant, is measured by

$$\omega = \frac{\phi - \phi_0}{t - t_0}. \tag{16}$$

If variable, it is measured by the limit of the same expression,

$\frac{d\phi}{dt} = \frac{\phi - \phi_0}{t - t_0}$, as the interval $t - t_0$ becomes indefinitely small.

The *angular acceleration* is the rate of change of angular velocity. If constant, it is measured by

$$\psi = \frac{\omega - \omega_0}{t - t_0}. \tag{17}$$

If variable, it is measured by the limit of the same expression, $\frac{d\omega}{dt} = \frac{\omega - \omega_0}{t - t_0}$, as the interval $t - t_0$ becomes indefinitely small.

If the radian be taken as the unit of angle, the dimensions of angle become

$$\left[\frac{\text{arc}}{\text{radius}}\right] = \frac{L}{L} = 1.$$

Hence the dimensions of angular velocity are T^{-1}, and of angular acceleration, T^{-2}.

If any point be revolving about a fixed point as a centre, its velocity in the circle varies as its angular velocity and the length of the radius jointly.

Angular velocities may be compounded by a process similar to that employed for the composition of motions.

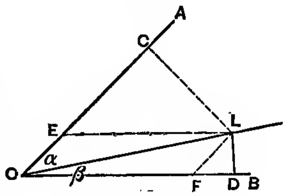

Fig. 16.

Let OA and OB (Fig. 16) represent two axes of rotation about which points are revolving with angular velocities ω_1 and ω_2 respectively; both rotations being clockwise when seen from the point O. The velocity at a point L, at unit distance from O, due to the motion about OA, is $\omega_1 \sin \alpha$, and that due to motion about OB is $\omega_2 \sin \beta$ in the opposite direction. The whole velocity of L is, therefore,

$$\omega_1 \sin \alpha - \omega_2 \sin \beta.$$

There must be some position of L for which this velocity becomes zero. Then $\omega_1 \sin \alpha = \omega_2 \sin \beta$. It follows at once that every point on the line OL is at rest. If we consider OL as the axis of rotation, and suppose the angular velocity of every point of the system about this axis to be ω, such that $\omega \sin \alpha = \omega_2 \sin (\alpha + \beta)$, this angular velocity will give the actual velocity of any point. To illustrate by a simple example, we will show that

$$\omega \sin \beta = \omega_2 \frac{\sin (\alpha + \beta)}{\sin \alpha} \sin \beta$$

is the velocity at B at unit distance from O. The velocity at B is only due to rotation about OA, and is therefore given by $\omega_1 \sin (\alpha + \beta)$. From our previous equation,

$$\omega_1 \sin \alpha = \omega_2 \sin \beta;$$

hence

$$\omega_1 \sin (\alpha + \beta) = \omega_2 \frac{\sin (\alpha + \beta)}{\sin \alpha} \sin \beta.$$

Now

$$\omega \sin \beta = \omega_2 \frac{\sin (\alpha + \beta)}{\sin \alpha} \sin \beta,$$

and the equality of the expressions is shown. Similarly it may be shown that the value of the velocity at any point N at unit distance from O, as given by the expression $\omega \sin NOL$, is equal to that given by $\omega_1 \sin AON - \omega_2 \sin BON$. The two rotations about OA and OB may thus be combined to form one rotation about OL.

Draw LF and LE parallel to OA and OB. Then the lines OE and OF are numerically proportional to the angular velocities ω_1 and ω_2; for, since $OELF$ is a parallelogram,

$$\frac{\sin\alpha}{\sin\beta} = \frac{OF}{OE}.$$

But, from our first equation,

$$\frac{\sin\alpha}{\sin\beta} = \frac{\omega_2}{\omega_1};$$

whence

$$OF : OE = \omega_2 : \omega_1.$$

The line OL is likewise proportional to ω, for, from the figure,

$$OF : OL = \sin\alpha : \sin(\alpha + \beta);$$

whence we see immediately, from the equation giving the value of ω, that

$$OL : OF = \omega : \omega_2.$$

We can therefore obtain the direction of the resultant axis, and the amount of the angular velocity, due to rotation about two other axes, by laying off on those axes, from their point of intersection, lengths numerically equal to the angular velocities about them, and drawing the diagonal of the parallelogram of which they are the sides. And so also any angular velocity may be resolved into three, the axes of which are at right angles to one another, by employing the trigonometrical functions of the angles which its axis makes with the three component axes.

It has been demonstrated that if a body be established in rotation, for any finite time, about an axis fixed with reference to points in it, however the position of the body be altered, it will continue to rotate with constant angular velocity about the same axis, unless constrained by outside forces to change its rotation. In other words, the axis of rotation always remains parallel to itself. This property is of importance in the discussion of some interesting applications of the preceding principles, which we shall next consider.

(1) The first of these is the method employed by Foucault to determine by experiment the fact of the earth's rotation. His apparatus consisted of a spherical pendulum bob, suspended by a truly cylindrical wire, so that it could swing freely in any plane. It can easily be seen, that, if such a pendulum were set up at the pole and swung, it would preserve its plane of oscillation invariable, and the earth would turn around under it, so that in twenty-four hours the pendulum would seem to have traversed a complete circle in the direction of the sun's apparent motion. At any other point on the earth's surface the change in apparent direction of the plane of oscillation would not be so great. Let ω represent the angular velocity of the earth, t the duration of the experiment, and ϕ the latitude. Let the pendulum be supposed to be at A (Fig. 17). Let NS be the earth's axis. Now, the angular velocity ω, represented by OC, may be resolved into two components, OD and DC, the axes of which lie respectively in the direction of the force acting on the pendulum and at right angles to it. The angular velocity DC has no influence in changing the relations of the pendulum and the earth; but the angular velocity $OD = OC \sin \phi = \omega \sin \phi$ is made evident by the rotation of a fixed line on the earth's sur-

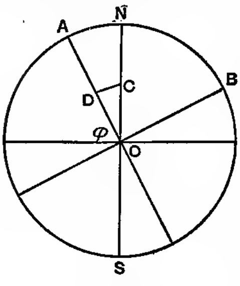

FIG. 17.

face, cutting the invariable plane of oscillation at the point of equilibrium of the pendulum. The plane of oscillation of the pendulum consequently appears to rotate in the opposite direction with an angular velocity $\omega \sin \phi$, and the angle swept out in any time t is $\omega t \sin \phi$. By such an apparatus has been determined, not only the fact of the earth's rotation, but even an approximate value of the length of the day.

(2) The phenomena presented by the gyroscope also offer an example of the application of the foregoing principles.

The construction of the apparatus can best be understood

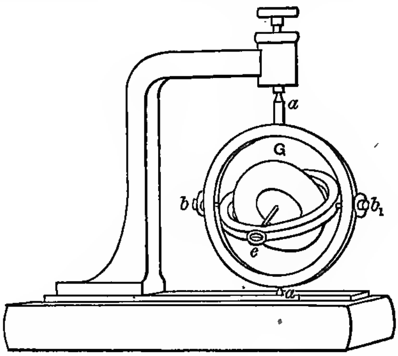

FIG. 18.

by the help of the diagram (Fig. 18). The outermost ring rests in a frame, and turns on the points a, a_1. The inner rests in the outer one, and turns on the pivots b, b_1, at right angles to the line of aa_1. Within this ring is mounted the wheel G, the axle of which is at right angles to the line bb_1, and in a plane passing through aa_1. At the point e is fixed a hook, from which weights may be hung. It is evident that if the wheel be mounted on the middle of the axle, the equilibrium of the apparatus is neutral in any position, and that a weight hung on the hook e will bring the axle of the wheel vertical, without moving the outer ring. If, however, the wheel be set in rapid rotation, with its axle horizontal, and a weight be hung on the hook, the whole system will revolve with a constant angular velocity about the points a, $a_{\prime\prime}$, and the axle of the wheel will remain horizontal.

The explanation of this phenomenon follows from the principles which we have already discussed. The conditions given are, that a body rotating with an angular velocity in one plane is acted on by a force tending to produce rotation in a perpendicular plane.

Let the plane of the paper represent the horizontal plane, and the line AB (Fig. 19) represent the direction of the axle at any moment. Lay off on OA a length OP proportional to the angular velocity of the wheel. If B be the point of application of the weight, the weight tends to turn the system about an axis CD at right angles to AB. Let us suppose, first, that, in the small interval of time t, the system acquires an angular velocity about CD proportional to OQ. Compounding the two angular velocities OP and OQ, we obtain the resultant OR. Now, resolving OQ parallel and at right angles to OR, we see that the parallel component is efficient in determining the length of OR, the component at right angles, the direction of OR. In the limit, as t becomes indefinitely small, OQ also becomes indefinitely small, and the resolved component Ox parallel to OR vanishes in comparison with OQ; because from the triangles we have $\frac{OQ}{OR}=\frac{Ox}{OQ}$. The effect will be a change of direction of the axle AB in the horizontal plane, without a change in the angular velocity of the wheel. This change is the equivalent of the introduction of a new angular velocity about an axis perpendicular to the plane of the paper. This new angular velocity, compounded with the angular velocity about OA, gives rise, as before, to a change

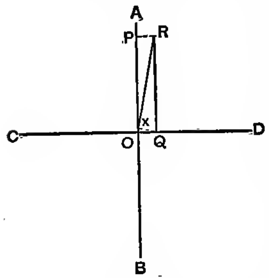

FIG. 19.

in the direction of the axis without a change in the angular velocity of the wheel; and this change in direction is such as to oppose the angular acceleration about CD, introduced by the weight at B. The system will revolve in a horizontal plane about O as a centre.

Another explanation, leading to the same results, has been given by Poggendorff. As has already been stated, it requires the application of a force to change the direction of the axis of a rotating body. This force is expended in changing the direction of motion of the component parts of the body. Poggendorf's explanation of the movements of the gyroscope is based on the action of couples formed by these separate forces.

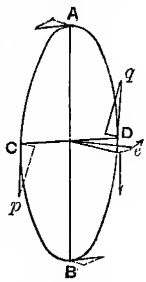

FIG. 20.

Let Fig. 20 represent the rotating wheel of the former diagram, the axle being supposed to be nearly horizontal. If the weight be hung at the point e, it tends to turn the wheel about a horizontal axis CD. The particles moving at A and at B in the plane CD offer no resistance to this change. Those at C moving downwards, and those at D moving upwards, act otherwise. The forces expressed by their momentum in the directions Cp and Dq are resolved into two each, one of them in the new plane assumed by the wheel, and the other at right angles to it. It will be seen that the latter component acts at C towards the right, and at D towards the left. There is thus set up a couple acting to turn the system about the axis AB counter-clockwise, as seen from A. As soon as this rotation begins, the particles moving at A out of the paper, and at B through the paper, are turned out of their original directions, and there arises another couple, of which the component at A is directed towards the left, and at B towards the right. This couple tends to cause the system to rotate about the axis CD counter-clockwise, as seen from C, and thus to oppose the tendency to rotation due to the weight at e.

All other points on the wheel except those in the lines AB and CD, are turned out of their paths by both rotations; and therefore components of the forces due to their motions appear in both couples in the final summation of effects. The result of the existence of these couples is a movement such as has already been described.

36. **Moment of Inertia.**—The *moment of inertia* of any body about an axis is defined as the summation of the products of the masses of the particles making up the body into the squares of their respective distances from the axis.

This product is the measure of the importance of the body's inertia with respect to rotation, and is proportional to the kinetic energy of the body having a given angular velocity about the axis; for, if any particle m, at a distance r from the axis, rotate with an angular velocity ω, its velocity is $r\omega$ and its kinetic energy is $\frac{1}{2}m\omega^2r^2$. The whole kinetic energy of the body is, therefore, $\frac{1}{2}\omega^2\Sigma mr^2$; and since we have assumed $\frac{1}{2}\omega^2$ to be constant, Σmr^2 is proportional to the kinetic energy of the rotating body. If we can find a distance k such that $\frac{1}{2}k^2\omega^2\Sigma m = \frac{1}{2}\omega^2\Sigma mr^2$, k is called the *radius of gyration*, and is the distance at which a mass equal to that of the whole body must be concentrated to possess the same moment of inertia as the body possesses.

The formula for moment of inertia is

$$I = \Sigma mr^2, \tag{18}$$

and its dimensions are ML^2.

The moment of inertia of a body with reference to an axis passing through its centre of inertia being known, its moment of inertia with reference to any other axis, parallel to this, is found by adding to the moment of inertia already known, the product of the mass of the body into the square of the distance of its centre of inertia from the new axis of rotation. For if the centre of inertia of the body of which we know the moment of inertia be C, and if m be any particle of that body, and if O be the new axis to which the moment of inertia is to be referred, making the construction as in Fig. 21, we have

$$r^2 = a^2 + 2r_{,}b + r_{,}^2.$$

Multiplying by the mass m, performing a similar operation for every particle of the body, and summing the results, we have

$$I_{,} = \Sigma ma^2 + \Sigma mr_{,}^2 + 2\Sigma mr_{,}b.$$

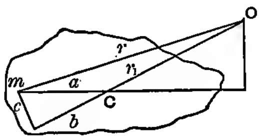

FIG. 21.

The term $2\Sigma mr_{,}b$ on the right vanishes, for we may write it $2r_{,}\Sigma mb$; and, since C is the centre of inertia, Σmb is zero (§ 33). Therefore

$$I_{,} = I + Mr_{,}^2. \tag{19}$$

This equation embodies the proposition which was to be proved.

The moment of inertia of the simple geometrical solids may be found by reckoning the moments of inertia for the separate particles of the body, and summing the results. We will show how this may be done in a few simple cases.

(1) To find the moment of inertia of a very thin rod AB, of length $2l'$ and mass $2m'$, about an axis xx', passing through the middle point:

Suppose the half-length to be divided into a very large number n of equal parts. The mass of each will be $\frac{m'}{n}$. The distance of the first from the axis is $\frac{l'}{n}$, of the second $\frac{2l'}{n}$, etc.

Their moments of inertia are

$$\frac{m'}{n} \times \frac{l'^2}{n^2}, \quad \frac{m'}{n} \times 4\frac{l'^2}{n^2} \ldots \frac{m'}{n} \times n^2\frac{l'^2}{n^2},$$

and the moment of inertia of the half-rod is

$$I' = \frac{m' l'^2}{n^3}(1 + 4 + 9 \ldots + n^2).$$

But $(1 + 4 + 9 \ldots + n^2)$, where n is indefinitely large, is $\dfrac{n^3}{3}$; hence $I' = \dfrac{m' l'^2}{3}$.

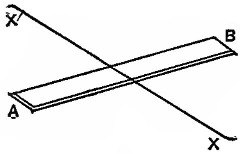

FIG. 22.

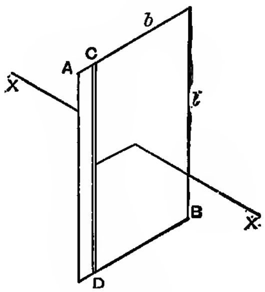

FIG. 23.

If l equal the whole length of the rod, m the whole mass, and I the entire moment of inertia,

$$I = \frac{ml^2}{12}. \tag{20}$$

(2) To find the moment of inertia of a thin plate AB (Fig. 23), of length l and breadth $2b'$, about an axis perpendicular to it and passing through its centre:

Suppose the half-plate to be divided into n rods, parallel to length: each rod will have a length l and a breadth $\frac{b'}{n}$. eir distances from the axis are $\frac{b'}{n}$, $\frac{2b'}{n}$, etc. Let m be the ss of the plate. The moment of inertia of each rod, with pect to an axis passing through its centre of inertia and pendicular to its length, is $\frac{m}{2n} \times \frac{l^2}{12}$. The moments of in- ia of the several rods about the parallel axis xx' are

$$\frac{m}{2n}\left(\frac{l^2}{12} + \frac{b'^2}{n^2}\right), \quad \frac{m}{2n}\left(\frac{l^2}{12} + 4\frac{b'^2}{n^2}\right), \text{ etc.,}$$

d the moment of inertia of the half-plate is

$$\frac{m}{2n} \times n\frac{l^2}{12} + \frac{m}{2n}\frac{b'^2}{n^2}(1 + 4 + 9 \ldots + n^2) = \frac{m}{2}\left(\frac{l^2}{12} + \frac{b'^2}{3}\right),$$

d of the whole plate equals

$$m\frac{l^2 + b^2}{12}. \tag{21}$$

A parallelopiped of which the axis is xx' may be supposed be made up of an infinite number of plates, such as AB. moment of inertia will be the moment of inertia of one ate multiplied by the number of plates; or, if M is the mass the parallelopiped, its moment of inertia is

$$\frac{M}{12}(l^2 + b^2). \tag{22}$$

The moment of inertia of any body, however irregular in form or density, may be found experimentally by the aid of another body of which the moment of inertia can be computed from its dimensions. We will anticipate the law of the pendulum, which has not been proved, for the sake of clearness. The body of which the moment of inertia is desired is set oscillating about an axis under the action of a constant force f. Its time of oscillation is, then,

$$t = \pi\sqrt{\frac{I}{f}},$$

where I is the moment of inertia.

If, now, another body, of which the moment of inertia can be calculated, be joined with the first, the time of oscillation alters to

$$t_{,} = \pi\sqrt{\frac{I + I_{,}}{f}},$$

where $I_{,}$ is the moment of inertia of the body added. Combining the two equations, we obtain, as the value of the moment of inertia desired,

$$I = \frac{I_{,}t^2}{t_{,}^2 - t^2}.$$

37. **Central Forces.**—If the velocity or direction of motion of a moving body in any way alter, we conceive it to be acted on by some force. In certain cases the direction of this force,

the law of its variation with the position of the body, may determined by considering the path or orbit traversed by body and the circumstances of its motion.

We shall illustrate this by a few propositions, selected on ount of their applicability in the establishment of the ory of universal gravitation. The proofs are substantially se given by Newton in the "Principia."

Proposition I.—If the radius vector, drawn from a fixed point a body moving in a curve, describe equal areas in equal es, the force which causes the body to move in the curve lirected towards the fixed point.

Let us suppose the whole time divided into equal periods, ing any one of which the body is not acted on by the force. vill, in the first period, move over a space represented by a ight line, as AB (Fig. 24). In the second period, it would, unhindered, move over an equal space BD and in the same :. Let us suppose it, however, deflected by a force acting tantaneously at the point B. It will move in a line BC such t, by hypothesis, triangle OBA = triangle OBC. Now, ngle ODB also = triangle OAB, therefore triangle OCB = ngle ODB, and CD is parallel to OB. Complete the parallogram $CDBE$; then it is evident that the motion BC is npounded of the motions BD and BE; and since forces are portional to the motions they asion, the force acting at B is portional to BE, and is directed ng the line BO. If now the iods into which the whole time

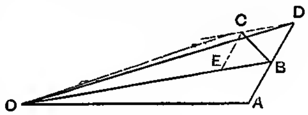

FIG 24.

divided become indefinitely small, in the limit the broken : ABC approaches indefinitely near to a curve, and the force ich causes the motion in the curve is always directed to the tre O.

Proposition II.—If a body move uniformly in a circle, the force acting upon it varies as its mass and the square of its velocity directly, and as the radius of the circle inversely.

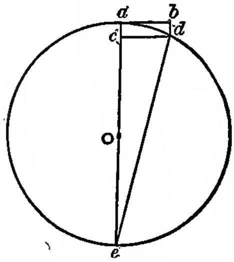

FIG. 25.

If the body m move in a circle (Fig. 25) with a constant angular velocity, and pass over, in any very small time t, the arc ad, which is so small that it may be taken equal to its chord, the motion may be resolved into two components ab and ac, one tangent and the other normal to the arc. Now f, the acceleration towards O, being constant for that small time, we have

$$s = ac = \tfrac{1}{2}ft^2.$$

The angle ade is a right angle, and therefore, by similar triangles, we have

$$ac = \frac{ad^2}{ae}.$$

But $ae = 2r$; ad represents the space traversed in the time t, and in the limit $\frac{ad}{t}$ represents v, the velocity in the circle. From the previous reasoning ac represents $\frac{1}{2}ft^2$; whence

$$\frac{ft^2}{2} = \frac{v^2t^2}{2r},$$

and

$$mf = \frac{mv^2}{r}.$$

Corollary I.—If two bodies revolve about the same centre, and the squares of their periodic times be in the same ratio as the cubes of the radii of their respective orbits, the forces ing on them will be inversely as the squares of their radii, l conversely. For, if T and $T_,$ represent the periodic times the two bodies moving in circles of radii r and $r_,$, with ve-ities v and $v_,$, then, by hypothesis,

$$T:T_, = \frac{2\pi r}{v} : \frac{2\pi r_,}{v_,} = r^{\frac{3}{2}} : r_,^{\frac{3}{2}};$$

ence

$$v : v_, = r_,^{\frac{1}{2}} : r^{\frac{1}{2}}.$$

w

$$f:f_, = \frac{v^2}{r} : \frac{v_,^2}{r_,};$$

ence

$$f:f_, = r_,^2 : r^2.$$

Corollary II.—The relation of Corollary I. holds with refer-ce to bodies describing similar parts of any similar figures ring the same centre. In the application of the proof, wever, we must substitute for uniform velocity the uniform scription of areas; and instead of radii we must use the tances of the bodies from the centre. The proof is as fol-vs:

If D and $D_,$ represent the radii of curvature of the paths of two bodies, R and $R_,$ the distances of the bodies from the itre of force, then, by hypothesis, letting A represent the a described in one period of time,

$$T : T_, = \frac{R^2}{A} : \frac{R_,^2}{A_,} = R^{\frac{3}{2}} : R_,^{\frac{3}{2}} = D^{\frac{3}{2}} : D_,^{\frac{3}{2}}$$

m the similarity of figures.

Now

$$A : A_, = vR : v_,R_, ;$$

hence

$$v : v_, = R_,^{\frac{1}{2}} : R^{\frac{1}{2}} = D_,^{\frac{1}{2}} : D^{\frac{1}{2}},$$

and

$$f : f_, = \frac{v^2}{D} : \frac{v_,^2}{D_,} = R_,^2 : R^2.$$

Proposition III.—If a body move in an ellipse, the force acting upon it, directed to the focus of the ellipse, varies inversely as the square of the radius vector.

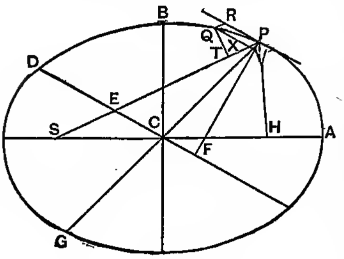

FIG. 26.

Suppose the body moving in the ellipse to be at the point P (Fig. 26), and the force to act upon it along the radius vector SP. At the point P draw the tangent PR, and from a point Q on the ellipse draw the chord Qv, cutting SP in x, and complete the parallelogram $PRQx$. From Q draw QT perpendicular to SP. Also draw the diameter GP and its conjugate DK. The force which acts on the body, causing it to leave the tangent PR and move in the line PQ, acts along SP, and in a time t (supposed very small) causes the body to move in the direction SP over the space Px; and since, in the small time considered, it may be assumed constant,

$$Px = \tfrac{1}{2}ft^2;$$

whence

$$t^2 = \frac{2Px}{f}.$$

Again: the area described by the radius vector in the time t is equal to $\frac{SP \cdot QT}{2}$; and if A represent the area described in unit time,

$$At = \frac{SP \cdot QT}{2}.$$

Equating these values of t, we obtain

$$\frac{SP \cdot QT^2}{4A^2} = \frac{2Px}{f};$$

whence

$$f = \frac{8A^2 \cdot Px}{QT^2} \cdot \frac{1}{SP^2}.$$

From Proposition I., the value of A is constant for any part of the ellipse. We shall now show that $\frac{Px}{QT^2}$ is also constant.

From similar triangles,

$$Px : Pv = PE : PC;$$

or, since by a property of the ellipse $PE = AC$,

$$Px : Pv = AC : PC.$$

Again, by another property of the ellipse,

$$Gv \cdot Pv : Qv^2 = PC^2 : CD^2.$$

If, now, we consider the time t to become indefinitely small, in the limit, P and Q approach indefinitely near; whence

$$Qv = Qx \quad \text{and} \quad Gv = 2PC.$$

The last proportion then becomes

$$PC \,.\, Pv : Qx^2 = PC^2 : 2CD^2.$$

Again, from similar triangles,

$$Qx : QT = PE : PF = AC : PF;$$

and from another property of the ellipse,

$$AC : PF = CD : CB;$$

whence

$$Qx : QT = CD : CB.$$

Combining these proportions,

$$\begin{aligned} Px : Pv &= AC : PC, \\ Pv : Qx^2 &= PC : 2CD^2, \\ Qx^2 : QT^2 &= CD^2 : CB^2, \end{aligned}$$

we obtain, finally,

$$Px : QT^2 = AC : 2CB^2;$$

that is, since AC and CB are constant, $\frac{Px}{QT^2}$ is constant.

We have now shown that, in the expression for the value of the force on the body at any point in the ellipse, all the factors are constant except $\frac{1}{SP^2}$. The force, therefore, varies inversely as the square of the radius vector.

CHAPTER II.

MASS ATTRACTION.

38. Mass Attraction.—The law of *mass attraction* was the first generalization of modern science. In its most complete form it may be stated as follows:—

Between every two material particles in the universe there is a stress, of the nature of an attraction, which varies directly as the product of the masses of the particles, and inversely as the square of the distance between them. This law is sometimes called the law of *universal attraction* and sometimes the law of *gravitation.*

Some of the ancient philosophers had a vague belief in the existence of an attraction between the particles of matter. This hypothesis, however, with the knowledge which they possessed, could not be proved. The geocentric theory of the planetary system, which obtained almost universal acceptance, offered none of those simple relations of the planetary motions upon which the law was finally established. It was not until the heliocentric theory, advocated by Copernicus, strengthened by the discoveries of Galileo, and systematized by the labors of Kepler, had been fully accepted, that the discovery of the law became possible.

In particular, the three laws of planetary motion published by Kepler in 1609 and 1619 laid the foundation for Newton's demonstrations. The laws are as follows:—

I. The planets move in ellipses of which one focus is situated at the sun.

II. The radius vector drawn from the sun to the planet sweeps out equal areas in equal times.

III. The squares of the periodic times of the planets are proportional to the cubes of their distances from the sun.

Kepler could give no physical reason for the existence of such laws. Later in the century, after Huyghens had discovered certain theorems relating to motion in a circle, it was seen that the third law would hold true for bodies moving in concentric circles, and attracted to the common centre by forces varying inversely as the squares of the radii of the circles. Several English philosophers, among them Hooke, Wren, and Halley, based a belief in the existence of an attraction between the sun and the planets upon this theorem.

The demonstration was by no means a rigorous one, and was not generally accepted. It was left for Newton to show that not only the third, but all, of Kepler's laws were completely satisfied by the assumption of the existence of an attraction acting between the sun and the planets, and varying inversely as the square of the distance. His propositions are substantially given in § 37.

Newton also showed that the attraction holding the moon in its orbit, which is presumably of the same nature as that existing between the sun and the planets, is of the same nature as that which causes heavy bodies to fall to the earth. This he accomplished by showing that the deviation of the moon from a rectilineal path is such as should occur if the force which at the earth's surface is the *force of gravity* were to extend outwards to the moon, and vary in intensity inversely as the square of the distance.

Two further steps were necessary before the final generalization could be reached. One was, to show the relation of the attraction to the masses of the attracting bodies; the other, to show that this attraction exists between all particles of matter, and not merely, as Huyghens believed, between those particles and the centres of the sun and planets.

The first step was taken by Newton. By means of pendulums having the same length, but with bobs of different materials, he showed that the force acting on a body at the earth's surface is proportional to the mass of the body, since all bodies have the same acceleration. He further brought forward, as the most satisfactory theory which he could form, the general statement that every particle of matter attracts and is attracted by every other particle.

The experiments necessary for a complete verification of this last statement were not carried out by Newton. They were performed in 1798 by Cavendish. His apparatus consisted essentially of a bar furnished at both ends with small leaden balls, suspended horizontally by a long fine wire, so that it turned freely in the horizontal plane. Two large leaden balls were mounted on a bar of the same length, which turned about a vertical axis coincident with the axis of rotation of the suspended bar. The large balls, therefore, could be set and clamped at any angular distance desired from the small balls. The whole arrangement was enclosed in a room, to prevent all disturbance. The motion of the suspended system was observed from without by means of a telescope. Neglecting as unessential the special methods of observation employed, it is sufficient to state that an attraction was observed between the large and small balls, and was found to be in accordance with the law as above stated.

39. **Measurement of the Force of Gravity.**—When two bodies attract one another, their relative motions are determined by Newton's third law. In the case of the attraction between the earth and a body near its surface, if we adopt a point on the earth's surface as the fixed point of reference, the acceleration of the body alone need be considered. Since the force acting upon it varies with its mass, and since its gain in momentum also varies with its mass, it follows that its acceleration will be constant, however its mass may vary. We may, therefore, obtain a direct measure of the

earth's attraction, or of the *force of gravity*, by allowing a body to fall freely, and determining its acceleration. It is found that a body so falling at latitude 40° will describe in one second about 16.08 feet, or 490 centimeters. Its acceleration is therefore 32.16 in feet and seconds or 980 in centimeters and seconds. We denote this acceleration by the symbol g.

The force acting on the body, or the *weight* of the body, is seen at once to be mg, where m is the mass of the body.

On account of the difficulties in the employment of this method, various others are used to obtain the value of g indirectly. For example, we may allow bodies to slide down a smooth inclined plane, and observe their motion. The force effective in producing motion on the plane is $g \sin \phi$, where ϕ is the angle of the plane with the horizontal; the space traversed in the time t is $s = \frac{1}{2}gt^2 \sin \phi$. By observing s and t, the value of g may be obtained. The motion is so much less rapid than that of a freely falling body that tolerably accurate observations can be made. Irregularities due to friction upon the plane and the resistance of the air, however, greatly vitiate any calculations based upon these observations. This method was used by Galileo, who was the first to obtain a measure of the acceleration due to the earth's attraction.

The most exact method for determining the value of g is based upon observations of the oscillation of a pendulum.

FIG. 27.

A *pendulum* may be defined as a heavy mass, or bob, suspended from a rigid support, so that it can oscillate about its position of equilibrium.

In the *simple*, or mathematical, pendulum, the bob is assumed to be a material particle, and to be suspended by a thread without weight. If the bob be stationary and acted on by gravity alone, the line of the thread will be the direction of the force. If the bob be withdrawn from the position of equilibrium (Fig. 27), it will be acted on by a force at right angles to the thread, in a direction opposite that of the displacement, expressed by $-g \sin \phi$, where ϕ is the angle between the perpendicular and the new position of the thread.

The force acting upon the bob at any point in the circle of which the thread is radius, if it be released, and allowed to swing in that circle, varies as the sine of the angle between the perpendicular and the radius drawn to that point. If we make the oscillation so small that the arc may be substituted for its sine without sensible error, the force acting on the bob varies as the displacement of the bob from the point of equilibrium.

A body acted on by a force varying as the displacement of the body from a fixed point will have a simple harmonic motion about its position of equilibrium.

Hence it follows that the oscillations of the pendulum are symmetrical about the position of equilibrium. The bob will have an amplitude on the one side of the vertical equal to that which it has on the other, and the oscillation, once set up, will continue forever unless modified by outside forces.

The importance of the pendulum as a means of determining the value of g consists in this: that, instead of observing the space traversed by the bob in one second, we may observe the number of oscillations made in any period of time, and determine the time of one oscillation; from this, and the length of the pendulum, we can calculate the value of g. The errors in the necessary observations and measurements are very slight in comparison with those of any other method.

40. Formula for Simple Pendulum.—The formula connecting the time of oscillation with the value of g is obtained as follows: The acceleration of the bob at any point in the arc is, as we have seen, $-g \sin \phi$, or $-g\phi$ if the arc be very small. The acceleration in a simple harmonic motion is

$$-\frac{4\pi^2}{T^2} a \cos \frac{2\pi t}{T}.$$

Since the bob has a simple harmonic motion, we may equate these expressions: hence

$$g\phi = \frac{4\pi^2}{T^2}a \cos\frac{2\pi t}{T}.$$

But $a \cos\frac{2\pi t}{T}$ is the displacement of the point having the simple harmonic motion, and is therefore equal to $l\phi$, if l represent the length of the thread: hence

$$g = \frac{4\pi^2 l}{T^2},$$

from which

$$T = 2\pi\sqrt{\frac{l}{g}}.$$

In this formula T represents the time of a double oscillation. It is customary to consider as a unit, the time of a single oscillation, when the formula becomes

$$t = \pi\sqrt{\frac{l}{g}}. \qquad (24)$$

41. **Physical Pendulum.**—Any pendulum fulfilling the requirements of the foregoing theory is, of course, unattainable in practice. We may, however, calculate, from the known dimensions and mass of the portions of matter making up the *physical pendulum*, what would be the length of a simple pendulum which would oscillate in the same time. It is clear that there must be some point in every physical pendulum the distance of which from the point of suspension is equal to the length of the corresponding simple pendulum; for the particles near the point of suspension tend to oscillate more rapidly than those

more remote, and the time of oscillation of the system, if it be rigid, will be intermediate between the times of oscillation which the particles nearest to, and most remote from, the point of suspension would have if they were oscillating freely. There will, therefore, be some one particle of which the proper rate of oscillation is the same as that of the whole pendulum. Its distance from the point of suspension is the length sought.

In determinations of the value of g by observations upon the time of oscillation of a pendulum, the length of the equivalent simple pendulum may be known in either of two ways.

(1) The pendulum may be constructed in such a manner that its moment of inertia and the position of its centre of gravity may be calculated. From these data the required length is readily obtained.

When the pendulum oscillates, each of its particles describes a simple harmonic motion, and passes through the mid-point of its path at the time that the pendulum passes through its position of equilibrium. The velocity of each particle at the mid-point of its path can therefore be expressed by $-\frac{2\pi a}{T}$, where T is the period of a complete oscillation, and a is an amplitude differing for each particle considered. Representing the distance of any particle from the axis of suspension by r, and the greatest value of the angular displacement of the pendulum by ϕ, we have $a = r\phi$. Hence the angular velocity of each particle, and therefore of the pendulum, is expressed by $-\frac{2\pi\phi}{T}$. The kinetic energy of a body, rotating about an axis with an angular velocity ω, has been shown in § 36 to be expressed by $\Sigma mr^2\frac{\omega^2}{2}$. Substituting in this expression the value obtained for the angular velocity of the pendulum, we obtain $\frac{1}{2}\Sigma mr^2\frac{4\pi^2\phi^2}{T^2}$ as the expression for the kinetic energy of the pendulum at the lowest point of its arc. At this point the pendulum possesses no potential energy. Its kinetic energy at this point must therefore be equal to its potential energy at the highest point of its arc, where it posesses no kinetic energy. If we represent by M the mass of the pendulum, and by R the distance of the centre of gravity from the point of suspension, $R\phi$ represents the distance traversed by the centre of gravity between the highest and the lowest points of its arc, and $\frac{1}{2}Mg\phi$ represents the average force acting on the centre of gravity between those points to produce rotation. The potential energy of the pendulum at the highest point of its arc is, therefore, $\frac{1}{2}MRg\phi^2$. Hence we have

$$\tfrac{1}{2}\Sigma mr^2\frac{4\pi^2\phi^2}{T^2} = \tfrac{1}{2}MRg\phi^2;$$

whence

$$T = 2\pi\sqrt{\frac{\Sigma mr^2}{MR\,.\,g}}. \tag{25}$$

This is the time of oscillation of a simple pendulum of which the length is $\frac{\Sigma mr^2}{MR}$. Therefore the moment of inertia of any physical pendulum divided by the product of its mass into the distance of its centre of gravity from the axis of suspension gives the length of the equivalent simple pendulum. An axis parallel to the axis of suspension, passing through the point on the line joining the axis of suspension with the centre of gravity of the pendulum and distant $\frac{\Sigma mr^2}{MR}$ from the axis of suspension, is called the *axis of oscillation.*

A pendulum consisting of a heavy spherical bob suspended by a cylindrical wire was used by Borda in his determinations of the value of g. The moment of inertia and the centre of vity of the system were easily calculated, and the length of simple pendulum to which the system was equivalent was s obtained.

(2) We may determine the length of the equivalent simple ıdulum directly by observation. The method depends upon principle that, if the axis of oscillation be taken as the s of suspension, the time of oscillation will not vary. The of of this principle is as follows:

Let r and $l-r$ represent the distances from the centre gravity to the axis of suspension and of oscillation rectively, m the mass of the pendulum, and I its moment inertia about its centre of gravity. Then, since the ment of inertia about the axis of suspension is $I+mr^2$, we 'e

$$l=\frac{I+mr^2}{mr}.$$

When the pendulum is reversed, we have

$$l_{,}=\frac{I+m(l-r)^2}{m(l-r)}.$$

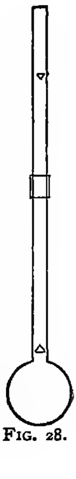

Fig. 28.

From the first equation we have $I=mr(l-r)$, which ue substituted in the second gives, after reduction, $=l$; that is, the length of the equivalent simple penum, and consequently the time of oscillation when pendulum swings about its axis of suspension, is the ıe as that when it is reversed, and swings about its mer axis of oscillation.

A pendulum (Fig. 28) so constructed as to take 'antage of this principle was used by Kater in his ermination of the value of g; and this form is known, :onsequence, as Kater's pendulum.

42. The Balance.—The comparison of masses is of such frequent occurrence in physical investigations that it is important to consider the theory of the balance and the methods of using it.

To be of value the balance must be accurate and sensitive; that is, it must be in the position of equilibrium when the scale-pans contain equal masses, and it must move out of that position on the addition to the mass in one pan of a very small fraction of the original load. These conditions are attained by the application of principles which have already been developed.

The balance consists essentially of a regularly formed beam, poised at the middle point of its length upon knife edges which rest on agate planes. From each end of the beam is hung a scale-pan in which the masses to be compared are

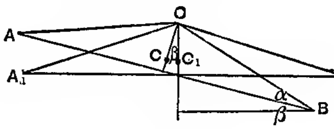

FIG. 29.

placed. Let O (Fig. 29) be the point of suspension of the beam; A, B, the points of suspension of the scale-pans; C, the centre of gravity of the beam, the weight of which is W. Represent $OA = OB$ by l, OC by d, and the angle OAB by α.

If the weight in the scale-pan at A be P, and that in the one at B be $P+p$, where p is a small additional weight, the beam will turn out of its original horizontal position, and assume a new one. Let the angle $COC_{,}$, through which it turns, be designated by β. Then the moments of force about O are equal; that is,

$$(P+p)l \,.\, \cos(\alpha+\beta) = Pl \,.\, \cos(\alpha-\beta) + Wd \,.\, \sin\beta\,;$$

from which we obtain, by expanding and transposing,

$$\tan\beta = \frac{pl \cos\alpha}{(2P+p)l \sin\alpha + Wd}. \qquad (26)$$

The conditions of greatest sensitiveness are readily deducible from this equation. So long as $\cos\alpha$ is less than unity, it is evident that $\tan\beta$, and therefore β, increases as the weight $2P$ of the load diminishes. As the angle α becomes less, the value of β also increases, until, when A, O, and B are in the same straight line, it depends only on $\frac{pl}{Wd}$, and is independent of the load. In this case $\tan\beta$ increases as d, the distance from the point of suspension to the centre of gravity of the beam, diminishes, and as the weight of the beam W diminishes. To secure sensitiveness, therefore, the beam must be as long and as light as is consistent with stiffness, the points of suspension of the beam and of the scale-pans must be very nearly in the same line, and the distance of the centre of gravity from the point of suspension of the beam must be as small as possible. Great length of beam, and near coincidence of the centre of gravity with the axis, are, however, inconsistent with rapidity of action. The purpose for which the balance is to be used must determine the extent to which these conditions of sensitiveness shall be carried.

Accuracy is secured by making the arms of the beam of equal length, and so that they will perfectly balance, and by attaching scale-pans of equal weight at equal distances from the centre of the beam.

In the balances usually employed in physical and chemical investigations, various means of adjustment are provided, by means of which all the required conditions may be secured. The beam is poised on knife edges; and the adjustment of its centre of gravity is made by changing the position of a nut which moves on a screw, placed vertically, directly above the point of suspension. Perfect equality in the moments of force due to the two arms of the beam is secured by a similar horizontal screw and nut placed at one end of the beam. The beam is a flat rhombus of brass, large portions of which are cut out so as to make it as light as possible. The knife edge on which the beam rests, and those upon which the scale-pans hang, are arranged so that, with a medium load, they are all nearly in the same line. A long pointer attached to the beam moves before a scale, and serves to indicate the deviation of the beam from the position of equilibrium. If the balance be accurately made and perfectly adjusted, and equal weights placed in the scale-pans, the pointer will remain at rest, or will oscillate through distances regularly diminishing on each side of the zero of the scale.

If the weight of a body is to be determined, it is placed in one scale-pan, and known weights are placed in the other until the balance is in equilibrium or nearly so. The final determination of the exact weight of the body is then made by one of three methods: we may continue to add very small weights until equilibrium is established; or we may observe the deviation of the pointer from the zero of the scale, and, by a table prepared empirically, determine the excess of one weight over the other; or we may place a known weight at such a point on a graduated bar attached to the beam that equilibrium is established, and find what its value is, in terms of weight placed in the scale-pan, by the relation between the length of the arm of the beam and the distance of the weight from the middle point of the beam.

If the balance be not accurately constructed, we can, nevertheless, obtain an accurate value of the weight desired. The method employed is known as Borda's method of double weighing. The body to be weighed is placed in one scale-pan, and balanced with fine shot or sand placed in the other. It is n replaced by known weights till equilibrium is again estab-.ed. It is manifest that the replacing weights represent weight of the body.

If the error of the balance consist in the unequal length of arms of the beam, the true weight of a body may be ob-ıed by weighing it first in one scale-pan and then in the ıer. The geometrical mean of the two values is the true ight; for let l_1 and l_2 represent the lengths of the two arms of balance, P the true weight, and P_1 and P_2 the values of the ights placed in the pans at the extremities of the arms of gths l_1 and l_2, which balance it. Then $Pl_2 = P_1l_1$ and $Pl_1 =$ $_2$; from which

$$P = \sqrt{P_1P_2}.$$

43. **Density of the Earth.**—One of the most interesting ›blems connected with the physical aspect of gravitation is determination of the density of the earth. It has been acked in several ways, each of which is worthy of consider-on.

The first successful determination of the earth's density s based upon experiments made in 1774 by Maskelyne. He served the deflection from the vertical of a plumb-line sus-ıded near the mountain Schehallien in Scotland. He then :ermined the density of the mountain by the specific gravity specimens of earth and rock from various parts of it, and culated the ratio of the volume of the mountain to that of earth. From these data the mean specific gravity of the ·th was determined to be about 4.7.

The next results were obtained from the experiments of vendish, in 1798, with the torsion balance already described. .e density, volume, and attraction of the leaden balls being own, the density of the earth could easily be obtained. The .ue obtained by Cavendish was about 5.5.

Another method, employed by Carlini in 1824, depends upon the use of the pendulum. The time of the oscillation of a pendulum at the sea-level being known, the pendulum is carried to the top of some high mountain, and its time of oscillation again observed. The value of g as deduced from this observation will, of course, be less than that obtained by the observation at the sea-level. It will not, however, be as much less as it would be if the change depended only on the increased distance from the centre of the earth. The discrepancy is due to the attraction of the mountain, which can, therefore, be calculated, and the calculations completed as in Maskelyne's experiment. The value obtained by Carlini by this method was about 4.8.

A fourth method, due to Airy, and employed by him in 1854, consists in observing the time of oscillation of a pendulum at the bottom of a deep mine. By § 29, (1), it appears that the attraction of a spherical shell of earth the thickness of which is the depth of the mine vanishes. The mean density of the earth may, therefore, be determined by the discrepancy between the values of g at the bottom of the mine and at the surface.

Still another method, used by Jolly, consists in determining by means of a delicate balance the increase in weight of a small mass of lead when a large leaden block is brought beneath it. Jolly's results were very consistent and give as the earth's density the value 5.69.

These methods have yielded results varying from that obtained by Airy, who stated the mean specific gravity to be 6.623, to that of Maskelyne, who obtained 4.7. The most elaborate experiments, by Cornu and Baille, by the method of Cavendish, gave as the value 5.56. This is probably not far from the truth.

When the density of the earth is known, we may calculate from it the value of the *constant of mass attraction*, that is, the attraction between two unit masses at unit distance apart.

Representing by D the earth's mean density, by R the earth's mean radius, and by k the constant of attraction, the mass of the earth is expressed by $\frac{4}{3}\pi R^3 D$. Since by § 29, (2), the attraction of a sphere is inversely as the square of the distance from its centre, the attraction of the earth on a gram at a point on its surface, or the weight of one gram, is expressed by $g = \frac{4}{3}\pi \frac{R^3 D}{R^2} k = \frac{4}{3}\pi R D k$. πR is twice the length of the earth's quadrant, or 2×10^9 centimetres. The value of g at latitude 40° is 980.11, and from the results of Cornu and Baille we may set D equal to 5.56. With these data we obtain k equal to 0.000000066 dynes.

44. **Projectiles.**—When a body is projected in any direction near the earth's surface, it follows, in general, a curved path. If the lines of force be considered as radiating from the earth's centre, this path will be, by Proposition III, § 37, an ellipse, with one focus at the earth's centre. If the path pursued be so small that the lines may be considered parallel, the centre of force is conceived of as removed to an infinite distance, and the curve becomes a parabola.

The fact that ordinary projectiles follow a parabolic path was first shown by Galileo, as a deduction from the principle which he established,—that a constant force produces a uniform acceleration. The proof is as follows: Suppose the body to be projected from the point O taken as origin, in the direction of the axis OY (Fig. 30), making any angle ϕ with OX, a vertical axis, and to move with a velocity $v = \frac{y}{t}$. Owing to the accelerating effect of gravity, it also moves in the vertical direction OX with a velocity $v_{,} = gt$. At any time t it will have traversed in the direction OY a space $y = vt$, and in the direction OX a space $x = \frac{1}{2}gt^2$. The co-ordinates of the position of the body

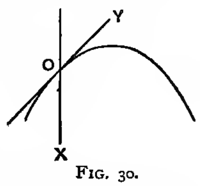

FIG. 30.

at any time t are, therefore, $y = vt$ and $x = \frac{1}{2}gt^2$. The equation connecting x and y becomes $y^2 = \frac{2v^2x}{g}$, which is the equation of a parabola referred to the diameter OX and the tangent OY. When the body is projected horizontally, the vertex of the parabola is at the origin of the motion. The body begins to approach the earth from the start, and reaches it at the same time that it would if allowed to fall freely.

One special case of importance in the consideration of the paths of projectiles is that in which the body moves in a circle. It is obvious, that, to bring about this result, the body must be projected horizontally with such an initial velocity that the acceleration due to the earth's attraction shall be precisely equal to the acceleration toward the centre which is necessary in order that the body should move in a circle (Proposition II, § 37). Hence we must have

$$\frac{mv^2}{R} = \frac{mM}{R^2}k,$$

where m and M are the masses of the body and the earth respectively, R is the earth's radius, and k the constant of attraction. Now v, the velocity of the body, equals

$$\frac{2\pi R}{T},$$

where T is the time of one complete revolution, and

$$M = \frac{4}{3}\pi R^3 D,$$

where D is the earth's mean density. Substituting these values, we obtain

$$\frac{4m\pi^2R^2}{T^2R} = \frac{4m\pi R^3D}{3R^2}k,$$

from which

$$T^2 = \frac{3\pi}{Dk}.$$

The result shows that the periodic time of any small body revolving about a sphere, and infinitely near its surface, is a function of the density only, and does not depend on the radius of the sphere.

Upon this principle Maxwell proposed, as an absolute unit of time, the time of revolution of a small satellite revolving infinitely near the surface of a globe of pure water at its maximum density.

CHAPTER III.

MOLECULAR MECHANICS.

CONSTITUTION OF MATTER.

45. **General Properties of Matter.**—Besides the properties already defined in § 3 as characteristic and essential, we find that all bodies possess the properties of compressibility and divisibility.

Compressibility.—All bodies change in volume by change of pressure and temperature. If a body of a given volume be subjected to pressure, it will return to its original volume when the pressure is removed, provided the pressure has not been too great. This property of assuming its original volume is called *elasticity*. The property of changing volume by the application of heat is sometimes specially called *dilatability*.

Divisibility.—Any body of sensible magnitude may, by mechanical means, be divided, and each of its parts may again be subdivided; and the process may be continued till the resulting particles become so minute that we are no longer able to recognize them, even when assisted by the most perfect appliances of the microscope. If the body be one that can be dissolved, it may be put in solution, and this may be greatly diluted; and in some cases the body may be detected by the color which it imparts to the diluent, even when constituting so small a proportion as one one-hundred-millionth part of the solution.

46. **Molecules.**—We are not, however, at liberty to conclude that matter is infinitely divisible. The fact, established

by observation, that bodies are impenetrable, and the one just noted, that they are also compressible, as well as other considerations, to be adduced later, lead to the opposite conclusion. To explain the coexistence of these properties, we are compelled to assume that bodies are composed of extremely small portions of matter, indivisible without destroying their identity, called *molecules*, and that these molecules are separated by interstitial spaces relatively larger, which are occupied by a highly elastic medium called the *ether*.

These molecules can be divided only by chemical means. The resulting subdivisions are called *atoms*. The atom, however, cannot exist in a free state. The molecule is the physical unit of matter, while the atom is the chemical unit.

47. **Composition of Bodies.**—It has just been said that atoms cannot exist in a free state. They are always combined with others, either of the same kind, forming simple substances, or of dissimilar kinds, forming compound substances.

There are about sixty-seven substances now known which cannot, in the present state of our knowledge, be decomposed, or made to yield anything simpler than themselves. We therefore call them *simple substances*, *elements*, or, if we desire to avoid expressing any theory concerning them, *radicals*. It is not improbable that some of these will yet be divided, perhaps all of them. We can call them elements, then, only provisionally.

48. **States of Aggregation.**—Bodies exist in three states, —the solid, the liquid, and the gaseous. In the *solid* state the form and volume of the body are both definite. In the *liquid* state the volume only is definite. In the *gaseous* state neither form nor volume is definite.

Many substances may, under proper conditions, assume either of these three states of aggregation; and some substances, as, for example, water, may exist in the three states under the same general conditions.

It is proper to add, however, that there is no such sharp line of distinction between the three states of matter as our definitions imply. Bodies present all gradations of aggregation between the extreme conditions of solid and gas; and the same substance, in passing from one state to the other, often presents all these gradations.

49. **Structure of Solids.**—With the exception of organized bodies, all solids may be divided into two classes. The bodies of one class are characterized by more or less regularity of form, which is called *crystalline;* those of the other class, exhibiting no such regularity, are called *amorphous.* For the formation of crystals a certain amount of freedom of motion of the molecules is necessary. Such freedom of motion is found in the gaseous and liquid states; and when crystallizable bodies pass slowly from these to the solid state, crystallization usually occurs. It may also occur in some solids spontaneously, or in consequence of agitation of the molecules by mechanical means, such as friction or percussion. Crystallizable bodies are called *crystalloids.*

Some amorphous bodies cannot, under any circumstances, assume the crystalline form. They are called *colloids.*

50. **Crystal Systems.**—Crystals are arranged by mineralogists in six systems.

In the first, or *Isometric,* system, all the forms are referred to three equal axes at right angles. The system includes the cube, the regular octahedron, and the rhombic dodecahedron.

In the second, or *Dimetric,* system, all the forms are referred to a system of three rectangular axes, of which only two are equal.

In the third, or *Hexagonal,* system, the forms are referred to four axes, of which three are equal, lie in one plane, and cross each other at angles of 60°. The fourth axis is at right angles to the plane of the other three, and passes through their common intersection.

The fourth, or *Orthorhombic*, system is characterized by three rectangular axes of unequal length.

In the fifth, or *Monoclinic*, system, the three axes are unequal. One of them is at right angles to the plane of the other two. The angles which these two make with each other, as well as the relative lengths of the axes, vary greatly for different substances.

In the sixth, or *Triclinic*, system, the three axes are oblique to each other, and unequal in length.

51. Forces determining the Structure of Bodies.—In view of what precedes, it is necessary to assume the existence of certain forces other than the mass attraction considered in § 38 acting between the molecules of matter. These forces seem to act only within very small or insensible distances, and vary with the character of the molecule. They are hence called *molecular forces*. In liquids and solids, there must be a force of the nature of attraction, holding the molecules together, and a force equivalent to repulsion, preventing actual contact. The attractive force is called *cohesion* when it unites molecules of the same kind, and *adhesion* when it unites molecules of different kinds. The repulsive force is probably a manifestation of that motion of the molecules which constitutes heat. In gases this motion is so great as to carry the molecules beyond the limit of their mutual molecular attractions: thus the apparent repulsion prevails, and the gas only ceases expanding when this repulsion is balanced by other forces.

52. Structure of the Molecule.—The facts brought to light in the study of crystals compel us to ascribe a structural form to the molecule, determining special points of application for the molecular forces. From this results the arrangement of molecules, which have the requisite freedom of motion, into regular crystalline forms.

FRICTION.

53. **General Statements.**—When the surface of one body is made to move over the surface of another, a resistance to the motion is set up. This resistance is said to be due to *friction* between the two bodies. It is most marked when the surfaces of two solids move over one another. It exists, however, also between the surfaces of a solid and of a liquid or a gas, and between the surfaces of contiguous liquids or gases. When the parts of a body move among themselves, there is a similar resistance to the motion, which is ascribed to friction among the molecules of the body. This internal friction is called *viscosity.*

54. **Laws of Friction.**—Owing to our ignorance of the arrangement and behavior of molecules, we cannot form a theory of friction based upon mechanical principles. The laws which have been found are almost entirely experimental, and are only approximately true even in the cases in which they apply.

It was found by Coulomb that, when one solid slides over another, the resistance to the motion is proportional to the pressure normal to the surfaces of contact, and is independent of the area of the surfaces and of the velocity with which the moving body slides over the other. It depends upon the nature of the bodies, and the character of the surfaces of contact. The ratio of the force required to keep the moving body in uniform motion to the force acting upon it normal to the surfaces of contact is called the *coefficient of friction.*

It was shown experimentally by Poiseuille that the rate of outflow of a liquid from a vessel through a long straight tube of very small diameter is proportional directly to the difference in pressure in the liquid at the two ends of the tube, to the fourth power of the radius of the tube, and inversely to the length of the tube. The flow of liquid under such conditions can be determined by mathematical analysis, and it is found

of being elevated, is depressed. If we change the tube for one of smaller bore, the water rises higher and the mercury sinks lower within it; but the rise or depression outside the tube remains the same. If we immerse the same tube in different liquids, we find that the heights to which they ascend vary for the different liquids. If, instead of changing the diameter, we change the thickness of the wall of the tube, no variation occurs in the amount of elevation or depression; and, finally, the rise or depression in the tube varies for any one liquid with its temperature.

57. **Law of Force assumed.**—It is found that a force such as is given by the law of mass attraction is not sufficient to produce these phenomena. They can, however, be explained if we assume an additional attraction between the molecules, such as we have already done. The expression, then, of the stress between two molecules m and m', at distance r, becomes

$$F = \frac{mm'}{r^2} + mm' f(r).$$

The only law which it is necessary to assign to the function of r in the second term is, that it is very great at insensible distances, diminishes rapidly as r increases, and vanishes while r, though measurable, is still a very small quantity. For adjacent molecules this molecular attraction is so much greater than the mass attraction, that it is customary, in the discussion of capillary phenomena, to omit the term $\frac{mm'}{r^2}$ from the expression for the force. The distance through which this attraction is appreciable is often called the *radius of molecular action*, and is denoted by the symbol ϵ. It is a very small distance, but is assumed to be much greater than the distance between adjacent molecules.

58. Methods of Development.—The different methods which have been employed to deduce, from this assumed attraction, results which could be submitted to experimental verification, are worthy of notice. They are distinct, though compatible with one another. Young was the first to treat the subject satisfactorily, though others had given partial and imperfect demonstrations before him. He showed that a liquid can be dealt with as if it were covered at the bounding surface with a stretched membrane, in which is a constant tension tending to contract it. From this basis he proceeded to deduce some of the most important of the experimental laws. Laplace, proceeding directly from the law of the attraction which we have already given, considered the attraction of a mass of liquid on a filament of the liquid terminating at the surface, and obtained an expression for the pressure within the mass at the interior end of the filament. He also was able, not only to account for already observed laws, but to predict, in at least one instance, a subsequently verified result. Some years later, Gauss, dissatisfied with Laplace's assumption, without *a priori* demonstration, of a known experimental fact, treated the subject from the basis of the principle of virtual velocities, which in this case is the equivalent of that of the conservation of energy. He proved, that, if any change be made in the form of a liquid mass, the work done or the energy recovered is proportional to the change of surface, and hence deduced a proof of the fact which Laplace assumed, and also an expression for the pressure within the mass of a liquid identical with his. For purposes of elementary treatment the earliest method is still the best. We shall accordingly employ the idea of *surface tension*, after having shown that it may be obtained from our first hypothesis.

59. Surface Tension.—Let us consider any liquid bounded by a plane surface, of which the line mn (Fig. 31) is the trace, and let the line $m'n'$ be the trace of a parallel plane at a distance ϵ from the plane of *mn*. The liquid is then divided into two parts by the plane of *m'n'*,—the general mass of the liquid, and a shell of thickness ϵ between the two planes. Then, if we imagine a plane passed through any point within the general mass, it is clear that the attraction of the molecules on opposite sides of that plane will give rise to a pressure normal to it, which will be constant for every direction of the plane; for the number of molecules now acting on the point is the same in all directions. Let, however, the point chosen be *P*, situated within the shell. With *P* as a centre, and with radius ϵ, describe a sphere. Now, it is evident that the number of mole-

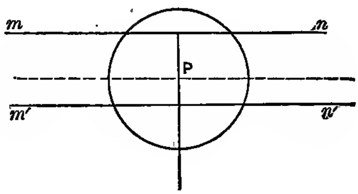

FIG. 31.

cules active in producing pressure upon the plane through *P*, parallel to *mn*, is less than that of those producing pressure upon the plane through *P* normal to *mn*. The pressure upon the parallel plane varies as we pass from the mass through the shell, from the value which it has within the mass, to zero, which it has at the plane *mn*. From this inequality of pressure in the two directions, parallel and normal to the surface, there results a stress or tension of the nature of a contraction in the surface.

Provided the radius of curvature of the surface be not very small, this tension will be constant for the surface of each liquid, or, more properly, for the surface of separation between two liquids, or a liquid and a gas.

60. Energy and Surface Tension.—We may here show how the energy of the liquid is related to the surface tension. It is plain, that, if the molecules, which by their mutual attractions give rise to the surface tension, be forced apart by the extrusion from the mass into the shell of a sheet of molecules along a plane normal to the surface, work will be done as the surface is increased. In every system free to move, movements will occur until the potential energy becomes a minimum: hence every free liquid moves so that its bounding surface becomes as small as possible; that is, it assumes a spherical form. This is exemplified in falling drops of water and in globules of mercury, and can be shown on a large scale by a method soon to be described. If we call the potential energy lost by a diminution in the surface of one unit, the *surface energy* per unit surface, we can show that it is numerically equal to the surface tension across one unit of length.

Suppose a thin film of liquid to be stretched on a frame

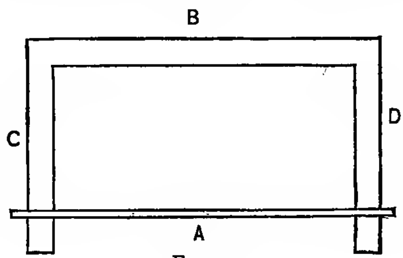

FIG. 32.

ABCD (Fig. 32), of which the part *BCD* is solid and fixed, and the part *A* is a light rod, free to slide along *C* and *D*. This film tends, as we have said, to diminish its free surface. As it contracts, it draws *A* towards *B*. If the length of *A* be *a*, and *A* be drawn towards *B* over *b* units, then if *E* represent the surface energy per unit of surface, the energy lost, or the work done, is expressed by *Eab*. If we consider the tension acting normal to A, the value of which is T for every unit of length, we have again for the work done during the movement of A, Tab. From these expressions we obtain at once $E = T$; that is, the numerical value of the surface energy per unit of surface is equal to that of the tension in the surface, normal to any line in it, per unit of length of that line.

61. Equation of Capillarity.—The surface tension introduces modifications in the pressure within the liquid mass (§ 85 *seq.*) depending upon the curvature of the surface. Consider any infinitesimal rectangle (Fig. 33) on the surface. Let the length of its sides be represented by s and $s_{\prime}$ respectively, and the radii of curvature of those sides by R and $R_{\prime}$. Also let ϕ and $\phi_{\prime}$ represent the angles in circular measure subtended by the sides from their respective centres of curvature. Now, a tension T for every unit of length acts normal to s and tangent to the surface. The total tension across s is then Ts; and if this tension be resolved parallel and normal to the normal at the point P, the centre of the rectangle, we obtain for the parallel component $Ts \sin \frac{\phi_{\prime}}{2}$, or, since $\phi_{\prime}$ is a very small angle, $Ts\frac{\phi_{\prime}}{2}$ or $Ts\frac{s}{2R_{\prime}}$. The opposite side gives a similar component; the side $s_{\prime}$ and the side opposite it give each a component $Ts_{\prime}\frac{s}{2R}$. The total force along the normal at P is then

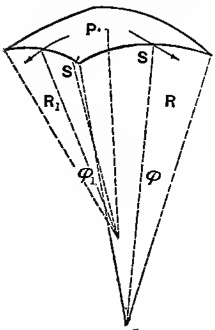

Fig. 33.

$$Tss_{\prime}\left(\frac{1}{R_{\prime}} + \frac{1}{R}\right);$$

and since $ss_{\prime}$ is the area of the infinitesimal rectangle, the force or pressure normal to the surface at P referred to unit of surface is

$$T\left(\frac{1}{R_{\prime}}+\frac{1}{R}\right).$$

From a theorem given by Euler we know that the sum $\frac{1}{R_{\prime}}+\frac{1}{R}$ is constant at any point for any position of the rectangular normal plane sections; hence the expression we have obtained fully represents the pressure at P.

If the surface be convex, the radii of curvature are positive, and the pressure is directed towards the liquid; if concave, they are negative, and the pressure is directed outwards. This pressure is to be added to the constant molecular pressure which we have already seen exists everywhere in the mass. If we denote this constant molecular pressure by K, the expression for the total pressure within the mass is

$$K+T\left(\frac{1}{R_{\prime}}+\frac{1}{R}\right),$$

where the convention with regard to the signs of $R_{\prime}$ and R must be understood. For a plane surface, the radii of curvature are infinite, and the pressure under such a surface reduces to K.

62. **Angles of Contact.**—Many of the capillary phenomena appear when different liquids, or liquids and solids, are brought in contact with one another. It becomes, therefore, necessary to know the relations of the surface tensions and the angles of contact. They are determined by the following considerations:

Consider first the case when three liquids meet along a line.

Let O represent the point where this line cuts a plane drawn at right angles to it. Then the tension T_{ab} of the surface of separation of the liquid a from the liquid b, acting normal to this line, is counterbalanced by the tensions T_{ac} and T_{bc} of the surfaces of separation of a and c, b and c. These tensions are always the same for the three liquids under similar conditions of temperature and purity. Knowing the value of the tensions, the angles which they make with one another are determined at once by the parallelogram of forces; and these angles are always constant.

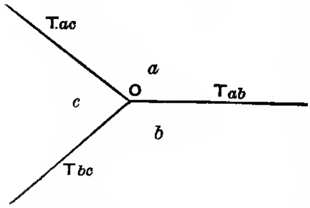

FIG. 34.

Similar relations arise if one of the liquids be replaced by a gas. Indeed, some experiments by Bosscha indicate that capillary phenomena occur at surfaces of separation between gases. We need, therefore, in the subsequent discussions, make no distinction between gases and liquids, and may use the general term fluids.

If T_{ab} be greater than the sum of T_{ac} and T_{bc}, the angle between T_{ac} and T_{bc} becomes zero, and the fluid c spreads itself out in a thin sheet between a and b. Thus, if a drop of oil be placed on water, the tension of the surface of separation between the air and water is greater than the sum of the tensions of the surfaces between the air and oil, and between the oil and water; hence the drop of oil spreads out over the water until it becomes almost indefinitely thin.

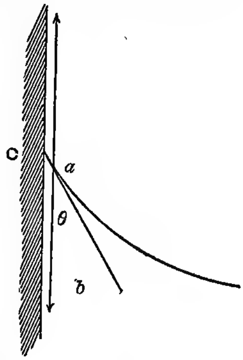

FIG. 35.

In the case of two fluids in contact with a plane solid (Fig. 35), it is evident that when the system is in equilibrium, the surface of separation between the fluids a and b, making the angle with the solid C, is

$$T_{ac} = T_{bc} + T_{ab} \cos \theta.$$

The angle of contact is then determined by the equation

$$\cos \theta = \frac{T_{ac} - T_{bc}}{T_{ab}}.$$

If T_{ac} be greater than $T_{ab} + T_{bc}$, the equation gives an impossible value for cos θ. In this case the angle becomes evanescent, the fluid b spreads itself out, and wets the whole surface of the solid. In other cases the value of θ is finite and constant for the same substances. Thus, a drop of water placed on a horizontal glass plate will spread itself over the whole plate; while a small quantity of mercury placed on the same plate will gather together into a drop, the edges of which make a constant angle with the surface.

63. Plateau's Experiments.—The preceding principles will enable us to explain a few of the most important experimental facts of capillarity.

A series of interesting results was obtained by Plateau from the examination of the behavior of a mass of liquid removed from the action of gravity. His method of procedure was to place a mass of oil in a mixture of alcohol and water, carefully mixed so as to have the same specific gravity as the oil. The oil then had no tendency to move as a mass, and was free to arrange itself entirely under the action of the molecular forces. Referring to the equation of Laplace, already obtained, it is evident that equilibrium can exist only when the sum $\left(\frac{1}{R} + \frac{1}{R}\right)$ is constant for every point on the surface. This is manifestly a property of the sphere, and is true of no other finite surface. Plateau found, accordingly, that the freely floating mass at once assumed a spherical form. This result we had previously reached by another method. If a solid body--for instance, a wire frame—be introduced into the mass of oil, of such a size as to reach the surface, the oil clings to it, and there is a break in the continuity of the surface at the points of contact. Each of the portions of the surface divided from the others by the solid then takes a form which fulfils the condition already laid down, that $\left(\frac{1}{R_{\prime}} + \frac{1}{R}\right)$ equals a constant. Plateau immersed a wire ring in the mass of oil. So long as the ring nowhere reached the surface, the mass remained spherical. On withdrawing a portion of the oil with a syringe, that which was left took the form of two equal *calottes*, or sections of spheres, forming a double convex lens. A mass of oil, filling a short, wide tube, projected from it at either end in a similar section of a sphere. As the oil was removed, the two end surfaces became less curved, then plane, and finally concave.

Plateau also obtained portions of other figures which fulfil the required condition. For example, a mass of oil was made to surround two rings placed at a short distance from one another. Portions of the oil were then gradually withdrawn, when two spherical *calottes* formed, one at each ring, and the mass between the rings became a right cylinder. It is evident that the cylinder fulfils the required condition for every point on its surface.

Plateau also studied the behavior of films. He devised a mixture of soap and glycerine, which formed very tough and durable films; and he experimented with them in air. Such films are so light that the action of gravity on them may be neglected in comparison with that of the surface tension. If the parts of the frame upon which the film is stretched be all in one plane, the film will manifestly lie in that plane. When, however, the frame is constructed so that its parts mark the edges of any geometrical volume, the films which are taken up by it often meet. Any three films thus meeting so arrange themselves as to make angles of 120° with one another. This follows as a consequence of the proposition which has already been given to determine the equilibrium of surfaces of separation meeting along a line. If four or more films meet, they always meet at a point.

Plateau also measured the pressure of air in a soap-bubble, and found that it differed from the external pressure by an amount which varied inversely as the radius of the bubble. This follows at once from Laplace's equation. This measurement also gives us a means of determining the surface tension; for, from Laplace's equation, the pressure inwards, due to the outer surface, is $T\frac{2}{R}$, and the pressure in the same direction due to the inner surface is also $T\frac{2}{R}$, for the film is so thin that we may neglect the difference in the radii of curvature of the two surfaces: hence the total pressure inwards is $\frac{4T}{R}$; and if this be measured by a manometer, we can obtain the value of T.

64. Liquids influenced by Gravity.—Passing now to consider liquid masses acted on by gravity, we shall treat only a few of the most important cases.

If a glass tube having a narrow bore be immersed perpendicularly in water, the water rises in the tube to a height inversely proportional to the diameter of the tube. This law is known as *Jurin's law.*

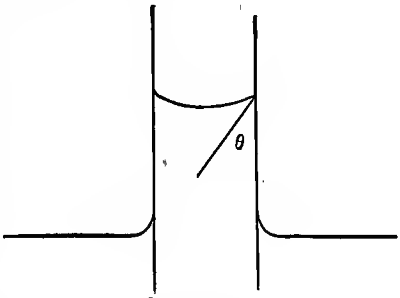

FIG. 36.

Let Fig. 36 represent the section of a tube of radius r immersed in a liquid, the surface of which makes an angle θ

with the wall. Then if T be the surface tension of the liquid, the tension acting upward is the component of this surface tension parallel to the wall, exerted all around the circumference of the tube. This is expressed by

$$2\pi r T \cos\theta.$$

This force, for each unit area of the tube, is

$$\frac{2\pi r T \cos\theta}{\pi r^2}.$$

The downward force, at the level of the free surface, making equilibrium with this, is due to the weight of the liquid column (§ 86). If we neglect the weight of the meniscus, this force per unit area, or the pressure, is expressed by hdg, where h is the height of the column and d the density of the liquid. We have, accordingly, since the column is in equilibrium,

$$\frac{2\pi r}{\pi r^2} T \cos\theta = hdg;$$

whence

$$h = \frac{2T\cos\theta}{rdg},$$

and the height is inversely as the radius of the tube.

If the liquid rise between two parallel plates of length l, separated by a distance r, the upward force per unit area is given by the expression $\frac{2l}{lr} T \cos\theta$, and the downward pressure by hdg; whence

$$h = \frac{2T\cos}{rdg},$$

and the height to which the liquid will rise between two such plates is equal to that to which it will rise in a tube the radius of which is equal to the distance between the plates.

If the two plates are inclined to one another so as to touch along one vertical edge, the elevated surface takes the form of a rectangular hyperbola; for let the line of contact of the plates be taken as the axis of ordinates, and a line drawn in the plane of the free surface of the liquid as the axis of abscissas, the elevation corresponding to each abscissa is inversely as the distance between the plates at that point, and the elevations are therefore inversely as the abscissas: hence the product of any abscissa by its corresponding ordinate is a constant. The extremities of the ordinates then mark out a rectangular hyperbola referred to its asymptotes.

65. Liquid Drops in Capillary Tubes.—When a drop of liquid is placed in a conical tube, it moves, if the surfaces are concave, towards the smaller end; if convex, towards the larger end. The explanation of these movements follows readily from the foregoing results. In case the surfaces are concave, letting θ (Fig. 37) be the angle of contact and α the angle of inclination of the wall of the tube to the axis, r and $r_{\prime}$ the radii of the tube at the extremities of the drop, r being the smaller of the two, then the expressions for the components of the tensions parallel with the axis acting in both cases outwards, are respectively

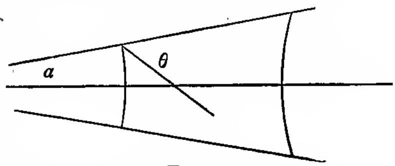

Fig. 37.

$$\frac{2\pi r T}{\pi r^2}\cos(\theta - \alpha),$$

and

$$\frac{2\pi r_{\prime} T}{\pi r_{\prime}^2}\cos(\theta + \alpha).$$

Of these two expressions the former is manifestly greater than the latter: hence the tendency of the drop is to move towards the smaller end of the tube.

If we assume that the concave surfaces are portions of spheres, of which R and $R_{,}$ are the respective radii of curvature, it follows that $r = R\cos(\theta - \alpha)$, and $r_{,} = R_{,}\cos(\theta + \alpha)$; hence the expressions for the tensions become $\frac{2T}{R}$ and $\frac{2T}{R_{,}}$. These are the values of the tensions as determined by Laplace's equation, and the movements of the drop might have been inferred directly from this equation by making the same assumption.

If a drop of water be introduced into a cylindrical capillary tube of glass, and if the air on the two ends of the drop have unequal pressures, the concavities thereby become unequal, the one on the side of the greater pressure presenting the greater concavity. The drop so circumstanced offers a resistance to this pressure; and it may, if the pressure be not too great, entirely counterbalance it. It is also evident, that, if several such drops be introduced successively, with intervening air-spaces, the pressure which they can unitedly sustain is equal to that which one can sustain multiplied by their number. Jamin found that, with a tube containing a large number of drops, a pressure of three atmospheres was maintained without diminution for fifteen days.

66. Movements of Solids.—In certain cases the action of the capillary forces produces movements in solid bodies partially immersed in a liquid. For example, if two plates, which are both either wetted or not wetted by the liquid, be partially immersed vertically, and brought so near together that the rise or depression of the liquid due to the capillary action begins, then the plates will move towards one another. In either case this movement is explained by the inequality of pressure on the two sides of each plate. When the liquid rises between the plates, the pressure is zero at that point in the column which lies in the same plane as the free external surface. At every internal point above this the molecules of the liquid are in a state of negative pressure or tension, and the plates are consequently drawn together. When the liquid is depressed between the plates, they are pressed together by the external liquid above the plane in which the top of the column between the plates lies. When one of the plates is wetted by the liquid and the other not, the plates move apart. This is explained by noting, that, if the plates be brought near together, the convex surface at the one will meet the concave surface at the other, and there will be a consequent diminution in both the elevation and the depression at the inner surfaces of the plates. The elevation and depression at the outer surfaces remaining unchanged, there will result a pull outwards on the wetted plate and a pressure outwards on the plate which is not wetted; and they will consequently move apart. Laplace showed, however, as the result of an extended discussion, that, though seeming repulsion exists between two plates such as we have just considered, yet, if the distance between the plates be diminished beyond a certain value, this repulsion changes to an attraction. This prediction has been completely verified by the most careful experiments.

67. Porous Bodies.—*Porous bodies* may be considered as assemblages of more or less irregular capillary tubes. Thus the explanation of many natural phenomena—as the wetting of a sponge, the rise of the oil in the wick of a lamp—follows directly from the preceding discussion.

DIFFUSION.

68. Solution and Absorption.—Many solid bodies, immersed in a liquid, after a while disappear as solids, and are taken up by the liquid. This process is called *solution*. The quantity of any body which a unit quantity of a given liquid will dissolve at a given temperature, is called its *solubility* in that liquid at that temperature. The solubility of a given solid varies greatly for different liquids, in many cases being so small as to be inappreciable.

Gases are also taken into solution by liquids. The process is usually called *absorption.* The quantity of gas dissolved in any liquid depends upon the temperature, and varies directly with the pressure. The solubility of any gas at a given temperature and at standard pressure is called its *coefficient of absorption* at that temperature.

Gases, in general, adhere strongly to the surfaces of solids with which they are in contact. This adhesion is so great, that the gases are sometimes condensed so as to form a dense layer which probably penetrates to some depth below the surface of the solid. The process is called the *absorption* of gases by solids. When the solid is porous, its exposed surface is greatly extended, and hence much larger quantities of gas are condensed on it than would otherwise be the case. When this condensation occurs there is in general a rise of temperature which may be so great as to raise the solid to incandescence. Thus, for example, spongy platinum, placed in a mixture of oxygen and hydrogen, becomes so heated as to inflame it.

69. **Free Diffusion of Liquids.**—When two liquids which are miscible are so brought together in a common vessel that the heavier is at the bottom and the lighter rests upon it in a well-defined layer, it is found that after a time, even though no agitation occur, they become uniformly mixed. Molecules of the heavier liquid make their way upwards through the lighter; while those of the lighter make their way downwards through the heavier, in apparent opposition to gravitation. *Diffusion* is the name which is employed to designate this phenomenon and others of a similar nature.

When one of the liquids is colored,—as, for example, solution of cupric sulphate,—while the other is colorless, the progress of the experiment may easily be watched and noted. When both liquids are colorless, small glass spheres, adjusted and sealed so as to have different but determinate specific gravities between those of the liquids employed, may be placed in the vessel used in the experiment, and will show by their positions the degree of diffusion which has occurred at any given time.

70. **Coefficient of Diffusion.**—Experiment shows that the amount of a salt in solution which at a given temperature passes, in unit time, through unit area of a horizontal surface, depends upon the nature of the salt and the rate of change of concentration at that surface,—that is, the quantity of a salt that passes a given horizontal plane in unit time is κCA, where A is the area, C the rate of change of concentration, and κ a coefficient that depends upon the nature of the substance. By *rate of change of concentration* is meant the difference in the quantities of salt in solution measured in grams per cubic centimetre, at two horizontal planes one centimetre apart, supposing the concentration to diminish uniformly from one to the other. It is plain, that, if C and A in the above expression be each equal to unity, the quantity of salt passing in unit time is κ. The quantity κ, called the *coefficient of diffusion*, is, therefore, the quantity of salt that passes in unit time through unit area of a horizontal plane when the difference of concentration is unity. Coefficients of diffusion increase with the temperature, and are found not to be entirely independent of the degree of concentration.

As implied above, the units of mass and length employed in these measurements are respectively the gram and the centimetre; but, since in most cases the quantity of salt that diffuses in one second is extremely small, it is usual to employ the day as the unit time.

71. Diffusion through Porous Bodies.—It was found by Graham that diffusion takes place through porous solids, such as unglazed earthenware or plaster, almost as though the liquids were in direct contact, and that a very considerable difference of pressure can be established between the two faces of the porous body while the rate of diffusion remains nearly constant.

72. Diffusion through Membranes.—If the membrane through which diffusion occurs be of a type represented by animal or vegetable tissue, the resulting phenomena, though in some respects similar, are subject to quite different laws. Colloid substances pass through the membrane very slowly, while crystalloid substances pass more freely. It is to be noted that the membrane is not a mere passive medium, as is the case with the porous substances already considered, but takes an active part in the process; and consequently one of the liquids frequently passes into the other more rapidly than would be the case if the surfaces of the liquids were directly in contact.

An explanation of these facts follows if we suppose that diffusion of a liquid through a continuous membrane can occur only when the liquid is capable of temporarily uniting with the membrane, and forming a part of it. Diffusion would then occur by the union of the liquid with the membrane on one face, and the setting free of an equal portion on the other.

If the membrane separate two crystalloids, it often happens that both substances pass through, but at different rates. In accordance with the usage of Dutrochet, we may say there is *endosmose* of the liquid, which passes more rapidly to the other liquid, and *exosmose* of the latter to the former. The whole process is frequently called *osmosis*. If the membrane be stretched over the end of a tube, into which the more rapid current sets, and the tube be placed in a vertical position, the liquid will rise in the tube until a very considerable pressure is attained. Dutrochet called such an instrument an *endosmometer*.

Graham made use of a similar instrument, which he called an *osmometer*, by means of which he studied, not only the action of porous substances, such as are mentioned above, but also that of various organic tissues; and he was able to reach quantitative results of great value. Pfeffer has more recently made an extended study of the phenomena of osmosis, especially in those aspects relating to physiological phenomena. He has shown that colloid membranes produced by purely chemical means are even more efficient than the organic membranes employed by Graham.

73. **Dialysis.**—Upon the principles just set forth Graham has founded a method of separating crystalloids from any colloid matters in which they may be contained, which is often of great importance in chemical investigations. The apparatus employed by Graham consists of a hoop, over one side of which parchment paper is stretched so as to constitute a shallow basin. In this basin is placed the mixture under investigation, and the basin is then floated upon pure water contained in an outer vessel. If crystalloids be present, they will in due time pass through the membrane into the water, leaving the colloids behind. The process is often employed in toxicology for separating poisons from ingesta or other matters suspected of containing them. It is called *dialysis*, and the substances that pass through are said to *dialyse*.

74. **Laws of Diffusion of Gases.**—Gases obey the same elementary laws of diffusion as liquids. The rate of diffusion varies inversely as the pressure, directly as the square of the absolute temperature, and inversely as the square root of the density of the gas. A gas diffuses through porous solids according to the same laws. An apparatus by which this may be conveniently illustrated consists of a porous cell, the open end of which is closed by a stopper, through which passes a

long tube. This is placed in a vertical position, with the open end of the tube in a vessel of water. If, now, a bell-jar containing hydrogen be placed over the porous cell, hydrogen passes into the cell more rapidly than the air escapes from it: the pressure inside is increased, as is shown by the escape of bubbles from the end of the tube. If, now, the jar be removed, diffusion outward occurs more rapidly than diffusion inward: the pressure within soon becomes less than the atmospheric pressure, as is shown by the rise of the water in the tube.

ELASTICITY.

75. Stress and Strain.—When a body is made the medium for the transmission of force, the application of Newton's third law shows that there is a stress in the medium. This stress is always accompanied by a corresponding change of form of the body, called a *strain.*

In some bodies equal stresses applied in any direction produce equal and similar strains. Such bodies are *isotropic.* In others the strain alters with the direction of the stress. These bodies are *eolotropic.*

According to the molecular theory of matter, the form of a body is permanent so long as the resultant of the stresses acting on it from without, with the interior forces existing between the individual molecules of the body, reduces to zero. The molecular forces and motions are such that there is a certain form of the body for every external stress in which its molecules are in equilibrium. Any change of the stress in the body is accompanied by a readjustment of the molecules, which is continued until equilibrium is again established.

If the stress tend only to increase or diminish the distance between the molecules, it is called a *tension* or a *pressure* respectively; if it tend to slide one line or sheet of molecules past another tangentially, it is called a *shear* or a *shearing-stress.*

All stresses can be resolved into these two forms. The corresponding changes of shape are called *dilatations*, *compressions*, and *shearing-strains*.

The term *pressure* is used with several different meanings. In order to most clearly present these, we will consider a right cylinder, transmitting a stress in the direction of its axis. The stress itself is often called the *total pressure* upon the cylinder.

If we conceive the cylinder to consist of a great number of elementary cylinders of small cross-section, and if the total pressure upon any one of them, as here defined, be to the total pressure on the whole cylinder as the cross-section of the elementary cylinder is to the cross-section of the whole cylinder, then it is said that the pressure on the cross-section is uniform, and the *pressure on an area* in that cross-section is defined as the product of the total pressure on the cylinder into the ratio of that area to the cross-section of the cylinder. Further, the *pressure at a point*, in a direction normal to the cross-section, is defined as the ratio of the pressure on an area, taken in the cross-section with its centre of inertia at the point, to that area, when the area is diminished indefinitely. This definition may at once be generalized. The pressure in any given direction at a point in a medium transmitting stress in any manner whatever, is the ratio of the pressure on any area, taken normal to the given direction and with its centre of inertia at the point, to that area, when the area is diminished indefinitely.

In case a stress exists between two bodies, which acts normally across a common surface of contact, the term pressure is also used to denote this stress, and the pressures on an area and at a point in the surface of contact, are defined exactly as above.

76. Modulus of Elasticity.—If, for a given amount of stress between certain limits, a body be deformed by a definite amount, which is constant so long as the stress remains constant, and if, when the stress is removed, the body regain its original condition, it is said to be *perfectly elastic.* Any body only partially fulfilling these conditions is said to be imperfectly elastic.

The definition of elasticity in its physical sense, as a property of bodies, has been already given. It is measured by the rate of change, in a unit of the body, of the stress with respect to the strain. Thus for example, the *voluminal elasticity* of a fluid is measured by the limit of the ratio of any small change of pressure to the corresponding change of unit volume. The *tractional elasticity* of a wire under tension is measured by the limit of the ratio of any small change in the stretching-weight to the corresponding change in unit length. This ratio is called the *modulus of elasticity*, or simply the *elasticity* of a body, and its reciprocal the *coefficient of elasticity.*

77. **Modulus of Voluminal Elasticity of Gases.**—Within certain limits of temperature and pressure the volume of any gas, at constant temperature, is inversely as the pressure upon it. This law was discovered by Boyle in 1662, and was afterwards fully proved by Mariotte. It is known, from its discoverer, as *Boyle's law.*

Thus, if p and $p_{\prime}$ represent different pressures, v and $v_{\prime}$ the corresponding volumes of any gas at constant temperature, then

$$p : p_{\prime} = \frac{1}{v} : \frac{1}{v_{\prime}};$$

whence

$$pv = p_{\prime}v_{\prime}. \tag{27}$$

Now, $p_{\prime}v_{\prime}$ is a constant which may be determined by choosing any pressure $p_{\prime}$ and the corresponding volume $v_{\prime}$ as standards: hence we may say, that, at any given temperature, the product pv is a constant. The limitations to this law will be noticed later.

If we draw the curve marked out by a point having its ordinate and abscissa so related that xy equals a constant, we obtain a rectangular hyperbola referred to its asymptotes. Let x represent the volume and y the pressure of a quantity of gas. Then this curve shows the relation of pressure and volume in all their combinations.

Draw the lines as in Fig. 38, letting AC, JD, represent volumes differing only by a small amount.

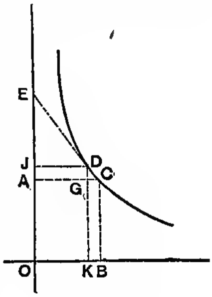

FIG. 38.

We must first show that AE is numerically equal to the modulus of elasticity. The ratio $\frac{CG}{AC}$ is the voluminal compression per unit volume for the increment of pressure GD: hence, by definition, $\frac{GD}{\frac{CG}{AC}}$ is the modulus of elasticity. But, from similarity of triangles, $AE : GD = AC : CG$.

Hence we have

$$AE = \frac{GD}{\frac{CG}{AC}} = \text{the modulus of elasticity.}$$

Now, since, by construction, the rectangles AB and JK are equal, and the rectangle AK is common to them, the rectangles JG and CK are equal, and

$$CG : DG = GA : GK.$$

By similar triangles,

$$CG : DG = CA : AE;$$

whence

$$GA : GK = CA : AE.$$

Now, if the increment of pressure be made indefinitely small, so that in the limit D and C coincide, the line CE becomes a tangent to the curve, and GA, GK, are respectively equal to CA, CB. CB therefore equals AE from the last proportion: hence, in the case of a gas obeying Boyle's law, the modulus of elasticity is numerically equal to the pressure.

The discussion of the experimental facts in connection with the elasticity of gases, and the explanation of the apparatus founded upon it, will be resumed in a future chapter.

78. Modulus of Voluminal Elasticity of Liquids.—When liquids are subjected to voluminal compression, it is found that their modulus of elasticity is much greater than that of gases. For at least a limited range of pressures the modulus of elasticity of any one liquid is constant, the change in volume being proportional to the change in the pressure. The modulus differs for different liquids.

The instrument used to determine the modulus of elasticity of liquids is called a *piezometer*. The first form in which the instrument was devised by Oersted, while not the best for accurate determinations, may yet serve as a type.

The liquid to be compressed is contained in a thin glass flask, the neck of which is a tube with a capillary bore. The flask is immersed in water contained in a strong glass vessel fitted with a water-tight metal cap, through which moves a piston. By the piston, pressure may be applied to the water, and through it to the flask and to the liquid contained in it.

The end of the neck of the small flask is inserted downwards under the surface of a quantity of mercury which lies at the bottom of the stout vessel. The pressure is registered by means of a compressed-air manometer (§ 96) also inserted in the vessel. When the apparatus is arranged, and the piston depressed, a rise of the mercury in the neck of the flask occurs, which indicates that the water has been compressed.

An error may arise in the use of this form of apparatus from the change in the capacity of the flask, due to the pressure. Oersted assumed, since the pressure on the interior and exterior walls was the same, that no change would occur. Poisson, however, showed that such a change would occur, and gave a formula by which it might be calculated. By introducing the proper corrections, Oersted's piezometer may be used with success.

A different form of the instrument, employed by Regnault, is, however, to be preferred. In it, by an arrangement of stopcocks, it is possible to apply the pressure upon either the interior or exterior wall of the flask separately, or upon both together, and in this way to experimentally determine the correction to be applied for the change in the capacity of the flask.

It is to be noted that the modulus of elasticity for liquids is so great, that, within the ordinary range of pressures, they may be regarded as incompressible. Thus, for example, the alteration of volume for sea-water by the addition of the pressure of one atmosphere is 0.000044. The change in volume, then, at a depth in the ocean of one kilometre, where the pressure is about 99.3 atmospheres, is 0.00437, or about $\frac{1}{230}$ of the whole volume.

79. Modulus of Voluminal Elasticity of Solids.—The modulus of voluminal elasticity of solids is believed to be generally greater than that of liquids, though no reliable experimental results have yet been obtained.

The modulus, as with liquids, differs for different bodies.

80. Shears.—A strain in which parallel planes or sheets of molecules are moved tangentially over one another, each plane being displaced by an amount proportional to its distance from one of the planes assumed as fixed, is called a *shear*.

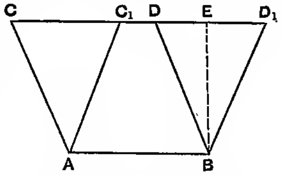

FIG. 39.

To illustrate this definition, let us consider a parallelopiped, of which the cross-section made at right angles to its sides is a rhombus, and let $ABDC$ in the diagram (Fig. 39) represent that cross-section.

If the rhombus $ABDC$ be deformed so as to become ABD_1C_1, that deformation is a simple shear. It is plain that a simple shear is equivalent to an extension in lines parallel to AD, and a contraction in those at right angles to AD. The directions AD and CB are called the *principal axes* of the shear. The *amount* of the shear is the displacement of the planes per unit of distance from the fixed plane; that is, $\frac{DD_1}{EB}$ is the amount of the shear.

The stresses that give rise to a simple shear can plainly be conceived of as consisting of two equal couples, the forces comprising which act tangentially upon parallel planes which are moved over one another, and make equal angles with the axes of the shear. The forces making up these couples may be compounded two and two, a and b, a_1 and b_1 (Fig. 40), making up a tension normal to the diminished axis; a_1 and b, a and b_1, making up a pressure normal to the increased axis. These stresses are measured per unit of area of the undeformed sides or sections of the solid.

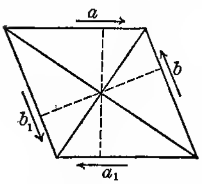

FIG. 40.

The resistance offered by a body to a shearing-stress is called its *rigidity*, and the ratio of a very small change in the stress to the corresponding increment in the amount of the shear is called the *modulus of rigidity*.

81. Elasticity of Tension.—The first experimental determinations of the relations between the elongation of a solid and the tension acting on it were made by Hooke in 1678. Experimenting with wires of different materials, he found that for small tensions the elongation is proportional to the stress. It was afterwards found that this law is true for small compressions.

The ratio of the stress to the elongation of unit length of a wire of unit section is the *modulus of tractional elasticity.* For different wires it is found that the elongation is proportional to the length of the wires, and inversely to their section. The formula embodying these facts is

$$e = \frac{Sl}{\mu s}, \tag{28}$$

where e is the elongation, l the length, s the section of the wire, S the stress, and μ the modulus of tractional elasticity.

A method of expressing the modulus of elasticity, due to Thomas Young, is sometimes valuable. "We may express the elasticity of any substance by the weight of a certain column of the same substance, which may be denominated the modulus of its elasticity, and of which the weight is such that any addition to it would increase it in the same proportion as the weight added would shorten, by its pressure, a portion of the substance of equal diameter." For example, considering a cubic litre of air at 0° C. and 760 millimetres of mercury pressure, and calling its weight unity, we find, from the fact that the weight of one litre of mercury is 10517 times that of a litre of air, that the pressure of the atmosphere upon a square decimetre is 79929 units. If we conceive the air as of equal density throughout, this pressure is equivalent to the weight of a column of air one square decimetre in section and 7992.9 metres high. The weight of this column is the modulus of elasticity for air; for we know, by Boyle's law, that if the column be altered in length, and its weight therefore correspondingly altered, the volume of the cubic litre of air under consideration will also alter inversely. The height of such a column of air as we have assumed is called the height of the *homogeneous atmosphere.*

82. **Elasticity of Torsion.**—When a cylindrical wire, clamped at one end, is subjected at the other to the action of a

couple the axis of which is the axis of the cylinder, it is found that the *amount of torsion*, measured by the angle of displacement of the arm of the couple, is proportional to the moment of the couple, to the length of the wire, and inversely to the fourth power of its radius. It also depends on the modulus of rigidity. The formulated statement of these facts is

$$\phi = \frac{Cl}{nr^4}, \tag{29}$$

where ϕ is the angle of torsion, l the length, r the radius of the wire, C the moment of couple, and n the modulus of rigidity. No general formula can be found for wires with sections of variable form.

The laws of torsion in wires were first investigated by Coulomb, who applied them in the construction of an apparatus of great value for the measurement of small forces.

The apparatus consists essentially of a small cylindrical wire, suspended firmly from the centre of a disk, upon which is cut a graduated circle. By the rotation of this disk any required amount of torsion may be given to the wire. On the other extremity of the wire is fixed, horizontally, a bar, to the ends of which the forces constituting the couple are applied. Arrangements are also made by which the angular deviation of this bar from the point of equilibrium may be determined. When forces are applied to the bar, it may be brought back to its former point of equilibrium by rotation of the upper disk. Let Θ represent the *moment of torsion;* that is, the couple which, acting on an arm of unit length, will give the wire an amount of torsion equal to a radian, C the moment of couple acting on the bar, τ the amount of torsion measured in radians; then

$$C = \Theta\tau.$$

We may find the value of Θ in absolute measure by a method of oscillations analogous to that used to determine g with the pendulum.

A body of which the moment of inertia can be determined by calculation is substituted for the bar, and the time T of one of its oscillations about the position of equilibrium observed.

Since the amount of torsion is proportional to the moment of couple, the oscillating body has a simple harmonic motion.

If a represent the amplitude of oscillation of any particle at distance r from the axis of rotation, we have $a = r\tau$. The velocity of the particle at the point of equilibrium is then

$$-\frac{2\pi a}{T},$$

and the angular velocity of the body, therefore, equals

$$-\frac{2\pi\tau}{T}.$$

The kinetic energy of a body rotating about a centre is $\frac{1}{2}I\omega^2$; and the kinetic energy of the body considered, at the point of equilibrium, is, therefore,

$$\tfrac{1}{2}I\frac{4\pi^2\tau^2}{T^2}.$$

The potential energy due to the torsion of the wire is $\frac{1}{2}\Theta\tau^2$, since $\frac{1}{2}\Theta\tau$ is the average moment of couple, and τ the distance through which this couple acts. These expressions are necessarily equal: hence

$$\tfrac{1}{2}I\frac{4\pi^2\tau^2}{T^2} = \tfrac{1}{2}\Theta\tau^2,$$

or

$$\Theta = \frac{4\pi^2 I}{T^2}.$$

We may use a single instead of a double oscillation, when we may write the formula

$$\Theta = \frac{\pi^2 I}{t^2}. \qquad (30)$$

This apparatus was used by Coulomb in his investigation of the law of electrical and magnetic actions. It was also employed by Cavendish, as has been already noticed, to determine the constant of gravitation.

83. Elasticity of Flexure.—If a rectangular bar be clamped by one end, and acted on at the other by a force normal to one of its sides, it will be bent or flexed. The amount of flexure—that is, the amount of displacement of the extremity of the bar from its original position—is found to be proportional to the force, to the cube of the length of the bar, and inversely to its breadth, to the cube of its thickness, and to the modulus of tractional elasticity. The formula therefore becomes

$$f = \frac{Fl^3}{\mu b d^3}. \qquad (31)$$

84. Limits of Elasticity.—The theoretical deductions and empirical formulas which we have hitherto been considering are strictly applicable only to perfectly elastic bodies. It is found that the voluminal elasticity of fluids is perfect, and that within certain limits of deformation, varying for different bodies, we may consider the elasticity of solids to be practically perfect for every kind of strain. If the strain be carried beyond the limit of perfect elasticity, the body is permanently deformed. This permanent deformation is called *set*.

Upon these facts we may base a distinction between solids and fluids: a *solid* requires the stress acting on it to exceed a certain limit before any permanent set occurs, and it makes no difference how long the stress acts provided it lie within the limits. A *fluid*, on the contrary, may be deformed by the slightest shearing stress, provided time enough be allowed for the movement to take place. The fundamental difference lies in the fact that fluids offer no resistance to shearing stress other than that due to internal friction or viscosity.

A solid, if it be deformed by a slight stress, is *soft;* if only by a great stress, is hard or *rigid.* A fluid, if deformed quickly by any stress, is *mobile;* if slowly, is *viscous.*

It must not be understood, however, that the behavior of elastic solids under stress is entirely independent of time. If, for example, a steel wire be stretched by a weight which is nearly, but not quite, sufficient to produce an immediate set, it is found that, after some time has elapsed, the wire acquires a permanent set. If, on the other hand, a weight be put upon the wire somewhat less than is required to break it, by allowing intervals of time to elapse between the successive additions of small weights, the total weight supported by the wire may be raised considerably above the *breaking-weight.* If the weight stretching the wire be removed, the return to its original form is not immediate, but gradual. If the wire carrying the weight be twisted, and the weight set oscillating by the torsion of the wire, it is found that the oscillations die away faster than can be explained by any imperfections in the elasticity of the wire.

These and similar phenomena are manifestly dependent upon peculiarities of molecular arrangement and motion. The last two are exhibitions of the so-called viscosity of solids. The molecules of solids, just as those of liquids, move among themselves, but with a certain amount of frictional resistance. This resistance causes the external work done by the body to be diminished, and the internal work done among the molecules becomes transformed into heat.

CHAPTER IV.

MECHANICS OF FLUIDS.

85. **Pascal's Law.**—A *perfect fluid* may be defined as a body which offers no resistance to shearing-stress. No actual fluids are perfect. Even those which approximate that condition most nearly, offer resistance to shearing-stress, due to their viscosity. With most, however, a very short time only is needed for this resistance to vanish; and all mobile fluids at rest can be dealt with as if they were perfect, in determining the conditions of equilibrium. If they are in motion, their viscosity becomes a more important factor.

As a consequence of this definition of a perfect fluid, follows a most important deduction. In a fluid in equilibrium, not acted on by any outside forces except the pressure of the containing vessel, the pressure at every point and in every direction is the same. This law was first stated by Pascal, and is known as *Pascal's law*.

The truth of Pascal's law appears, if, in a fluid fulfilling the conditions indicated, we imagine a cube of the fluid to become solidified. Then, if the law as just stated were not true, there would be an unbalanced force in some direction, and the cube would move, which is contrary to the statement that the fluid is in equilibrium. If a vessel filled with a fluid be fitted with a number of pistons of equal area A, and a force Ap be applied to one of them, acting inwards, a pressure Ap will act outwards upon the face of each of the pistons. These pressures may be balanced by a force applied to each piston. If $n + 1$ be the number of the pistons, the outward pressure on n of them, caused by the force applied to one, is npA.

The fluid will be in equilibrium when a pressure p is acting on unit area of each piston. It is plain that the same reasoning will hold if the area of one of the pistons be A and of another be nA. A pressure Ap on the one will balance a pressure of nAp on the other. This principle governs the action of the *hydrostatic press*.

86. Relations of Fluid Pressures due to Outside Forces.—If forces, such as gravitation, act on the mass of a fluid from without, Pascal's law no longer holds true. For suppose the cube of solidified fluid to be acted on by gravity; then the pressure on the upper face must be less than that on the lower face by the weight of the cube, in order that the fluid may still be in equilibrium. As the cube may be made as small as we please, it appears that, in the limit, the pressure on the two faces only differs by an infinitesimal; that is, the pressure in a fluid acted on by outside forces is the same at one point for all directions, but varies continuously for different points.

The surface of a fluid of uniform density acted on by gravity, if at rest, is everywhere perpendicular to the lines of force; for, if this were not so, the force at a point on the surface could be resolved into two components, one normal and the other tangent to the surface. But, from the nature of a fluid, the tangential force would set up a motion of the fluid, which is contrary to the statement that the fluid is at rest. If a surface be drawn through the points in the field at which the pressure is the same, that surface will be perpendicular to the lines of force. For, consider a filament of solidified fluid lying in the surface; its two ends suffer equal and opposite pressures; hence, since by hypothesis the fluid is in equilibrium, the force acting upon it, due to gravity, can have no component in the direction of its length, and is perpendicular to the surface in which it lies.

Surfaces of equal pressures are equipotential surfaces. In small masses of fluid, in which the lines of force due to gravity are parallel, these surfaces are horizontal planes. In larger masses, such as the oceans, they are curved to correspond to the divergence of the lines of force from the centre of the earth.

In a liquid the pressure at a point is proportional to its depth below the surface of the liquid. For, imagine two rectangular prisms of solidified liquid with bases which are equal and coincident with the surface of the liquid, and with heights such that the one is n times the other. From the fact that liquids are practically incompressible, the weight of these prisms acting downwards is proportional to their volumes, and hence to their heights. Since the liquid is in equilibrium, these weights are balanced by the upward pressures on their lower bases. These pressures are therefore proportional to the heights of the prisms, or to the depths of the surfaces to which they are applied.

From the foregoing principles, it is evident that a liquid contained in two communicating vessels of any shape whatever will stand at the same level in both. If one, however, be filled with a liquid of different density from that in the other, equilibrium will be established when the depths are inversely as the densities of the liquids.

87. **The Barometer.**—The instrument best adapted to illustrate these principles, and also of great importance in many physical investigations, is the *barometer*. It was invented by Torricelli, a pupil of Galileo. The fact that water can be raised in a tube in which a complete or partial vacuum has been made was known to the ancients, and was explained by them, and by the schoolmen after them, by the maxim that "Nature abhors a vacuum." They must have been familiar with the action of pumps, for the force-pump, a far more complicated instrument, was invented by Ctesibius of Alexandria, who lived during the second century B.C. It was not until the time of Galileo, however, that the first recorded observations were made that

the column of water in a pump rises only to a height of about 10.5 metres. Galileo failed to give the true explanation of this fact. He had, however, taught that the air has weight; and his pupil Torricelli, using that principle, was more successful.

He showed, that if a glass tube sealed at one end, over 760 millimetres long, were filled with mercury, the open end stopped with the finger, the tube inverted, and the unsealed end plunged beneath a surface of mercury in a basin, on withdrawing the finger the mercury in the tube sank until its top surface was about 760 millimetres above the surface of the mercury in the basin. The specific gravity of the mercury being 13.59, the weight of the mercury column and that of the water column in the pump agreed so nearly as to show that the maintenance of the columns in both cases was due to a common cause,—the pressure of the atmosphere. This conclusion was subsequently verified and established by Pascal, who requested a friend to observe the height of the mercury column at the bottom and at the top of a mountain. On making the observation, the height of the column at the top was found to be less than at the bottom. Pascal himself afterwards observed a slight though distinct diminution in the height of the column on ascending the tower of St. Jacques de la Boucherie in Paris.

The form of barometer first made by Torricelli is still often used, especially when the instrument is stationary, and is intended to be one of precision. In the finest instruments of this class a tube is used which is three or four centimetres in diameter, so as to avoid the correction for capillarity. A screw of known length, pointed at both ends, is arranged so as to move vertically above the surface of the mercury in the cistern. When an observation is to be made, the screw is moved until its lower point just touches the surface. The distance between its upper point and the top of the column is measured by means of a cathetometer; and this distance added to the length of the screw gives the height of the column.

Other forms of the instrument are used, most of which are arranged with reference to convenient transportability. Various contrivances are added by means of which the column is made to move an index, and thus record the pressure on a graduated scale. All these forms are only modifications of Torricelli's original instrument.

The pressure indicated by the barometer is usually stated in terms of the height of the column. Mercury being practically incompressible, this height is manifestly proportional to the pressure at any point in the surface of the mercury in the cistern. The pressure on any given area in that surface can be calculated if we know the value of g at the place and the specific gravity of mercury, as well as the height of the column. The standard barometric pressure, represented by 760 millimetres of mercury, is a pressure of 1.033 kilograms on every square centimetre. It is called a pressure of one *atmosphere;* and pressures are often measured by atmospheres.

In the preparation of an accurate barometer, it is necessary that all air be removed from the mercury: otherwise it will collect in the upper part of the tube, by its pressure lower the top of the column, and make the barometer read too low. The air is removed by partially filling the tube with mercury, which is then boiled in the tube, gradually adding small quantities of mercury, and boiling after each addition, until the tube is filled. The boiling must not be carried too far; for there is danger, in this process, of expelling the air so completely that the mercury will adhere to the sides of the tube, and will not move freely. For rough work the tube may be filled with cold mercury, and the air removed by gently tapping the tube, so inclining it that the small bubbles of air which form can coalesce, and finally be set free at the surface of the mercury.

88. Archimedes' Principle.—If a solid be immersed in a fluid, it loses in weight an amount equal to the weight of the fluid displaced. This law is known, from its discoverer, as *Archimedes' principle.*

The truth of this law will appear if we consider the space occupied by the solid as filled with the fluid. The fluid in this space will then be in equilibrium, and the upward pressure on it must exceed the downward pressure by an amount equal to its weight. The resultant of the pressure acts through the centre of gravity of the assumed portion of fluid, otherwise equilibrium would not exist. If, now, the solid occupy the space, the difference between the upward and the downward pressures on it must still be the same as before,—namely, the weight of the fluid displaced by the solid; that is, the solid loses in apparent weight an amount equal to the weight of the displaced fluid.

89. **Floating Bodies.**—When the solid floats on the fluid, the weight of the solid is balanced by the upward pressure. In order that the solid shall be in equilibrium, these forces must act in the same line. The resultant of the pressure, which lies in the vertical line passing through the centre of gravity of the displaced fluid, must pass through the centre of gravity of the solid. Draw the line in the solid joining these two centres, and call it the *axis* of the solid. The equilibrium is stable when, for any infinitesimal inclination of the axis from the vertical, the vertical line of upward pressure cuts the axis in a point above the centre of gravity of the solid. This point is called the *metacentre.*

90. **Specific Gravity.**—Archimedes' principle is used to determine the specific gravity of bodies. The *specific gravity* of a body is defined as the ratio of its weight to the weight of an equal volume of pure water at a standard temperature.

The specific gravity of a solid that is not acted on by water may be determined by means of the *hydrostatic balance.* The body under examination, if it will sink in water, is suspended from one scale-pan of a balance by a fine thread, and is weighed.

It is then immersed in water, and is weighed again. The difference between the weights in air and in water is the weight of the displaced water, and the ratio of the weight of the body to the weight of the displaced water is the specific gravity of the body.

If the body will not sink in water, a sinker of unknown weight and specific gravity is suspended from the balance, and counterpoised in water. Then the body, the specific gravity of which is sought, is attached to the sinker, and it is found that the equilibrium is destroyed. To restore it, weights must be added to the same side. These, being added to the weight of the body, represent the weight of the water displaced.

The specific gravity of a liquid is obtained by first balancing in air a mass of some solid, such as platinum or glass, that is not acted on chemically by the liquid, and then immersing the mass successively in the liquid to be tested and in water. The ratio of the weights which must be used to restore equilibrium in each case is the specific gravity of the liquid.

The specific gravity of a liquid may also be found by means of the *specific gravity bottle*. This is a bottle fitted with a ground glass stopper. The weight of the water which completely fills it is determined once for all. When the specific gravity of any liquid is desired, the bottle is filled with the liquid, and the weight of the liquid determined. The ratio of this weight to the weight of an equal volume of water is the specific gravity of the liquid.

The same bottle may be used to determine the specific gravity of any solid which cannot be obtained in continuous masses, but is friable or granular. A weighed amount of the solid is introduced into the bottle, which is then filled with water, and the weight of the joint contents of the bottle determined. The difference between the last weight and the sum of the weights of the solid and of the water filling the bottle is the weight of the water displaced by the solid. The ratio of the weight of the solid to the weight thus obtained is the specific gravity of the solid.

The specific gravity of a liquid may also be obtained by means of *hydrometers*. These are of two kinds,—the hydrometers of *constant weight* and those of *constant volume*. The first consists usually of a glass bulb surmounted by a cylindrical stem. The bulb is weighted, so as to sink in pure water to some definite point on the stem. This point is taken as the zero; and, by successive trials with different liquids of known specific gravity, points are found on the stem to which the hydrometer sinks in these liquids. With these as a basis, the divisions of the scale are determined and cut on the stem.

The hydrometer of constant volume consists of a bulb weighted so as to stand upright in the liquid, bearing on the top of a narrow stem a small pan, in which weights may be placed. The weight of the hydrometer being known, it is immersed in water; and, by the addition of weights in the pan, a fixed point on the stem is brought to coincide with the surface of the water. The instrument is then transferred to the liquid to be tested, and the weights in the pan changed until the fixed point again comes to the surface of the liquid. The sum of the weight of the hydrometer and the weights added in each case gives the weight of equal volumes of water and of the liquid, from which the specific gravity sought is easily obtained.

The specific gravity of gases is often referred to air or to hydrogen instead of water. It is best determined by filling a large glass flask, of known weight, with the gas, the specific gravity of which is to be obtained, and weighing it, noting the temperature and the pressure of the gas in the flask. The weight of the gas at the standard temperature and pressure is then calculated, and the ratio of this weight to the weight of the same volume of the standard gas is the specific gravity desired. The weight of the flask used in this experiment must be very exactly determined. The presence of the air vitiates all weighings performed in it, by diminishing the true weight of the body to be weighed and of the weights employed, by an amount proportional to their volumes. The consequent error is avoided either by performing the weighings in a vacuum produced by the air-pump, or by correcting the apparent weight in air to the true weight. Knowing the specific gravity of the weights and of the body to be weighed, and the specific gravity of air, this can easily be done.

91. Motions of Fluids.—If the parts of the fluid be moving relatively to each other or to its bounding-surface, the circumstances of the motion can be determined only by making limitations which are not actually found in nature. There thus arise certain definitions to which we assume that the fluid under consideration conforms.

The motion of a fluid is said to be *uniform* when each element of it has the same velocity at all points of its path. The motion is *steady* when, at any one point, the velocity and direction of motion of the elements successively arriving at that point remain the same for each element. If either the velocity or direction of motion change for successive elements, the motion is said to be *varying*. The motion is further said to be *rotational* or *irrotational* according as the elements of the fluid have or have not an angular velocity about their axes.

In all discussions of the motions of fluids a condition is supposed to hold, called the *condition of continuity*. It is assumed that, in any volume selected in the fluid, the change of density in that volume depends solely on the difference between the amounts of fluid flowing into and out of that volume. In an incompressible fluid, or liquid, if the influx be reckoned plus and the efflux minus, we have, letting Q represent the amount of the liquid passing through the boundary in any one direction, $\Sigma Q = 0$. The results obtained in the discussion of fluid motions must all be interpreted consistently with this condition. If the motion be such that the fluid breaks up into discontinuous parts, any results obtained by hydrodynamical considerations no longer hold true.

If we consider any stream of incompressible fluid, of which the cross-sections at two points where the velocities of the elements are v_1 and v_2 have respectively the areas A_1 and A_2, we can deduce at once from the condition of continuity

$$A_1v_1 = A_2v_2. \qquad (32)$$

92. **Velocity of Efflux.**—We shall now apply this principle to discover the *velocity of efflux* of a liquid from an orifice in the walls of a vessel.

Consider any small portion of the liquid, bounded by stream lines, which we may call a *filament*. Represent the velocity of the filament at B (Fig. 41) by v_1, and at C by v, and the areas of the cross-sections of the elements at the same points by A_1 and A. We have then, as above, $A_1v_1 = Av$.

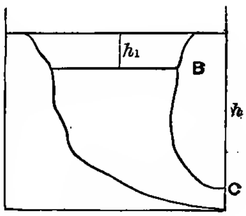

FIG. 41.

We assume that the flow has been established for a time sufficiently long for the motion to become steady. The energy of the mass contained in the filament between B and C is, therefore, constant. Let V_1 represent the potential at B due to gravity, V the potential at C, and d the density of the liquid. The mass that enters at B in a unit of time is

$$dA_1v_1.$$

The mass that goes out at C is the equal quantity dAv. The energy entering at B is

$$dA_1v_1(\tfrac{1}{2}v_1^2 + V_1),$$

the energy passing out at C is

$$dAv(\tfrac{1}{2}v^2 + V).$$

If the pressures at B and C on unit areas be expressed by p_1 and p, the work done at B on the entering mass by the pressure p_1 is $p_1A_1v_1$, and at C on the outgoing mass is pAv. The energy within the filament remaining constant, the incoming must equal the outgoing energy; therefore

$$pAv + dAv(\tfrac{1}{2}v^2 + V) = p_1A_1v_1 + dA_1v_1(\tfrac{1}{2}v_1^2 + V_1),$$

whence, since $A_1v_1 = Av$, we have

$$\frac{p}{d} + \tfrac{1}{2}v^2 + V = \frac{p_1}{d} + \tfrac{1}{2}v_1^2 + V_1.$$

We may write this equation

$$\tfrac{1}{2}(v^2 - v_1^2) = (V_1 - V) + \frac{p_1 - p}{d}; \qquad (33)$$

or, again, since $A_1v_1 = Av$,

$$\tfrac{1}{2}v^2\left(1 - \frac{A^2}{A_1^2}\right) = (V_1 - V) + \frac{p_1 - p}{d}. \qquad (34)$$

To apply equation (34) to the case of a liquid flowing freely into air from an orifice at C, we observe that the difference of potential $(V_1 - V)$ equals the work done in carrying a gram from C to B or equals $g(h - h_1)$, where h represents the height of the surface above C, and h_1 that of the surface above B. Further, we have

$$p_1 = p_a + dgh_1,$$

where p_a is the atmospheric pressure. At the orifice p equals p_a. We have then

$$\tfrac{1}{2}v^2\left(1 - \frac{A^2}{A_1^2}\right) = g(h - h_1) + gh_1 = gh,$$

whence

$$v^2 = \frac{2gh}{1 - \dfrac{A^2}{A_1^2}}.$$

If, now, A become indefinitely small as compared with A_1, in the limit, the velocity at C becomes

$$v = \sqrt{2gh}; \tag{35}$$

that is, the velocity of efflux of a small stream issuing from an orifice in the wall of a vessel is independent of the density of the liquid, and is equal to the velocity which a body would acquire in falling freely through a distance equal to that between the surface of the liquid and the orifice.

This theorem was first given by Torricelli from considerations based on experiment, and is known as *Torricelli's theorem.*

We may apply the general equation to the case of the efflux of a liquid through a siphon. A ***siphon*** is a bent tube which is used to convey a liquid by its own weight over a barrier. One end of the siphon is immersed in the liquid, and the discharging end, which must be below the level of the liquid, opens on the other side of the barrier. To set the siphon in operation it must be first filled with the liquid, after which a steady flow commences.

In this case, as before, we may set $\frac{A^2}{A_1^2} = 0$, $v_1 = 0$, p and p_1 both $= p_a$, and $(V_2 - V) = gl$, where l is the distance between the surface level and the discharging orifice. The velocity becomes $v = \sqrt{2gl}$. The siphon, therefore, discharges more rapidly the greater the distance between the surface level and the orifice. It is manifest that the height of the bend in the tube cannot be greater than that at which atmospheric pressure would support the liquid.

The flow of a liquid into the vacuum formed in the tube of an ordinary pump may also be discussed by the same equation. The *pump* consists essentially of a tube, fitted near the bottom with a partition, in which is a valve opening upwards. In the tube slides a tightly fitting piston, in which is a valve, also opening upwards. The piston is first driven down to the partition in the tube, and the enclosed air escapes through the valve in the piston. When the piston is raised, the liquid in which the lower end of the tube is immersed passes through the valve in the partition, rises in the tube and fills the space left behind the piston. When the piston is again lowered, the space above it is filled with the liquid, which is lifted out of the tube at the next up-stroke.

To determine the velocity of the liquid following the piston, we notice that in this case $p_1 = p_a$ and $p = 0$ if the piston move upward very rapidly, $(V_1 - V) = -gh$, where h is the height of the top of the liquid column above the free surface in the reservoir, and $\frac{A^2}{A_1^2}$ again $= 0$. We then have

$$\tfrac{1}{2}v^2 = \frac{p_a}{d} - gh.$$

The velocity when $h = 0$ is

$$v = \sqrt{\frac{2p_a}{d}}.$$

When h is such that $dgh = p_a$, $v = 0$, which expresses the condition of equilibrium.

The equation $v = \sqrt{\frac{2p_a}{d}}$ expresses, more generally, the velocity of efflux, through a small orifice, of any fluid of density d, from a region in which it is under a constant pressure p_a, into a vacuum.

Torricelli's theorem is shown to be approximately true by allowing liquids to run from an orifice in the side of a vessel, and measuring the path of the stream. If the theorem be true, this ought to be a parabola, of which the intersection of the plane of the stream and of the surface of the liquid is the directrix; for each portion of the liquid, after it has passed the orifice, will behave as a solid body, and move in a parabolic path. The equation of this path is found, as in § 44, to be

$$\frac{2v^2}{g} x = y^2.$$

Now by Torricelli's theorem, we may substitute for v^2 its value $2gh$, whence $y^2 = 4hx$. In this equation, since the initial movement of the stream is supposed to be horizontal, the perpendicular line through the orifice being the axis of the parabola, and the orifice being the origin, h is the distance from the orifice to the directrix. Experiments of this kind have been frequently tried, and the results found to approximate more nearly to the theoretical as various causes of error were removed.

When, however, we attempt to calculate the amount of liquid discharged in a given time, there is found to be a wider discrepancy between the results of calculation and the observed facts. Newton first noticed that the diameter of the jet at a short distance from the orifice is less than that of the orifice. He showed this to be a consequence of the freedom of motion among the particles in the vessel. The particles flow from all directions towards the orifice, those moving from the sides necessarily issuing in streams inclined towards the axis of the jet. Newton showed that by taking the diameter of the narrow part of the jet, which is called the *vena contracta*, as the diameter of the orifice, the calculated amount of liquid escaping agreed far more closely with theory.

When the orifice is fitted with a short cylindrical tube, the interference of the different particles of the liquid is in some degree lessened, and the quantity discharged increases nearly to that required by theory.

93. Diminution of Pressure.—The Sprengel air-pump, an important piece of apparatus to be described hereafter, depends for its operation on the diminution of pressure at points along the line of a flowing column of liquid. Let us consider a large reservoir filled with liquid, which runs from it by a vertical tube entering the bottom of the reservoir. From Eq. 34 the value of p, the pressure at any point in the tube, is

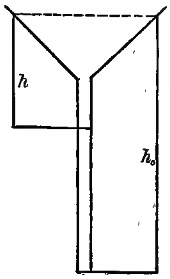

Fig. 42.

$$p = p_1 + (V_1 - V)d - \tfrac{1}{2}dv^2\left(1 - \frac{A^2}{A_1^2}\right).$$

The ratio $\dfrac{A^2}{A_1^2}$ may be set equal to zero. If h (Fig. 42) represent the height of the upper surface above the point in the tube at which we desire to find the pressure, then $(V_1 - V) = gh$. We then have

$p = p_a + dgh - \frac{1}{2}dv^2$. If the tube be always filled with the liquid, $Av = A_0v_0$, where A and A_0 represent the areas of the cross-sections of the tube at the point we are considering and at the bottom of the tube, and v and v_0 represent the corresponding velocities. Further, $v_0^2 = 2gh_0$ if h_0 represent the distance from the upper surface to the bottom of the tube. We obtain, by substitution,

$$p = p_a + dg\left(h - \frac{A_0^2}{A^2}h_0\right). \qquad (36)$$

If h equal $\frac{A_0^2}{A^2}h_0$, we have $p = p_a$; and if an opening be made in the wall of the tube, the moving liquid and the air will be in equilibrium. If h be less than $\frac{A_0^2}{A^2}h_0$, the pressure p will be less than p_a, and air will flow into the tube. Since this inequality exists when $A_0 = A$, it follows, that, if a liquid flow from a reservoir down a cylindrical tube, the pressure at any point in the wall of the tube is less than the atmospheric pressure by an amount equal to the pressure of a column of the liquid, the height of which is equal to the distance between the point considered and the bottom of the tube.

94. **Vortices.**—A series of most interesting results has been obtained by Helmholtz, Thomson, and others, from the discussion of the rotational motions of fluids. Though the proofs are of such a nature that they cannot be presented here, the results are so important that they will be briefly stated.

A *vortex line* is defined as the line which coincides at every point with the instantaneous axis of rotation of the fluid element at that point. A *vortex filament* is any portion of the fluid bounded by vortex lines.

A *vortex* is a vortex filament which has "contiguous to it over its whole boundary irrotationally moving fluid."

The theorems relating to this form of motion, as first proved by Helmholtz, in 1868, show that,—

(1) A vortex in a perfect fluid always contains the same fluid elements, no matter what its motion through the surrounding fluid may be.

(2) The *strength* of a vortex, which is the product of its angular velocity by its cross-section, is constant; therefore the vortex in an infinite fluid must always be a closed curve, which, however, may be knotted and twisted in any way whatever.

(3) In a finite fluid the vortex may be open, its two ends terminating in the surface of the fluid.

(4) The irrotationally moving fluid around a vortex has a motion due to its presence, and transmits the influence of the motion of one vortex to another.

(5) If the vortices considered be infinitely long and rectilinear, any one of them, if alone in the fluid, will remain fixed in position.

(6) If two such vortices be present parallel to one another, they revolve about their common centre of gravity.

(7) If the vortices be circular, any one of them, if alone, moves with a constant velocity along its axis, at right angles to the plane of the circle, in the direction of the motion of the fluid rotating on the inner surface of the ring.

(8) The fluid encircled by the ring moves along its axis in the direction of the motion of the ring, and with a greater velocity.

(9) If two circular vortices move along the same axis, one following the other, the one in the rear moves faster, and diminishes in diameter; the one in advance moves slower, and increases in diameter. If the strength and size of the two be nearly equal, the one in the rear overtakes the other, and passes through it. The two now having changed places, the action is repeated indefinitely.

(10) If two circular vortices of equal strength move along

the same axis toward one another, the velocities of both gradually decrease and their diameters increase. The same result follows if one such vortex move toward a solid barrier.

The preceding statements apply only to vortices set up in a perfect fluid. They may, however, be illustrated by experiment. To produce circular vortices in the air, we use a box which has one of its ends flexible. A circular opening is cut in the opposite end. The box is filled with smoke or with finely divided sal-ammoniac, resulting from the combination of the vapors of ammonia and hydrochloric acid. On striking the flexible end of the box, smoke rings are at once sent out.

The smoke ring is easily seen to be made up of particles revolving about a central core in the form of a ring. With such rings many of the preceding statements may be verified.

An illustration of the open vortex is seen when an oar-blade is drawn through the water. By making such open vortices, using a circular disk, many of the observations with the smoke-rings may be repeated in another form.

95. **Air-Pumps.**—The fact that gases, unlike liquids, are easily compressed, and obey Boyle's law under ordinary conditions of temperature and pressure, underlies the construction and operation of several pieces of apparatus employed in physical investigations. The most important of these is the *air-pump*.

The working portion of the air-pump is constructed essentially like the common lifting-pump already described. The valves must be light and accurately fitted. The vessel from which the air is to be exhausted is joined to the pump by a tube, the orifice of which is closed by the valve in the bottom of the cylinder.

A special form of vessel much used in connection with the air-pump is called the *receiver*. It is usually a glass cylinder, open at one end, and closed by a hemispherical portion at the other. The edge of the cylinder at the open end is ground perfectly true, so that all points in it are in the same plane. This ground edge fits upon a plane surface of roughened brass, or ground glass, called the *plate*, through which enters the tube which joins the receiver to the cylinder of the pump. The joint between the receiver and the plate is made tight by a little oil or vaseline.

The action of the pump is as follows: as the piston is raised, the pressure on the upper surface of the valve in the cylinder is diminished, and the air in the vessel expands in accordance with Boyle's law, lifts the valve, and distributes itself in the cylinder, so that the pressure at all points in the vessel and the cylinder is the same. The piston is now forced down, the lower valve is closed by the increased pressure on its upper surface, the valve in the piston is opened, and the air in the cylinder escapes. At each successive stroke of the pump this process is repeated, until the pressure of the remnant of air left in the vessel is no longer sufficient to lift the valves.

The density of the air left in the vessel after a given number of strokes is determined, provided there be no leakage, by the relations of the volumes of the vessel and the cylinder.

Let V represent the volume of the vessel, and C that of the cylinder when the piston is raised to the full extent of the stroke. Let d and d_1 respectively represent the density of the air in the vessel before and after one stroke has been made. After one down and one up stroke have been made, the air which filled the volume V now fills $V+C$. It follows that

$$\frac{d_1}{d} = \frac{V}{V+C}.$$

As this ratio is constant no matter what density may be considered, it follows that, if d_n represent the density after n strokes,

$$\frac{d_n}{d} = \left(\frac{V}{V+C}\right)^n. \tag{37}$$

As this fraction cannot vanish until n becomes infinite, it is plain that a perfect vacuum can never, even theoretically, be obtained by means of the air-pump. If, however, the cylinder be large, the fraction decreases rapidly, and a few strokes are sufficient to bring the density to such a point that either the pressure is insufficient to lift the valves, or the leakage through the various joints of the pump counterbalances the effect of longer pumping.

In the best air-pumps the valves are made to open auto-

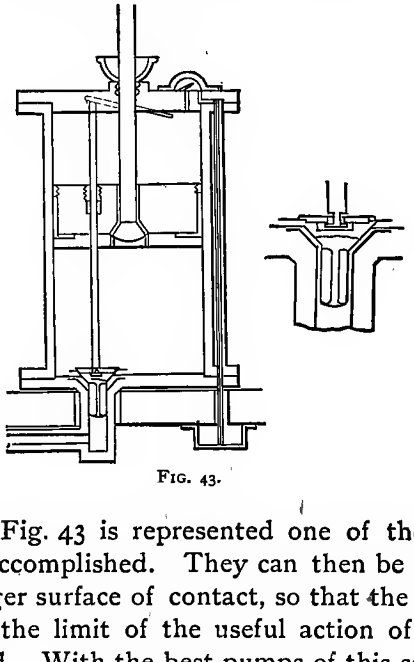

FIG. 43.

matically. In Fig. 43 is represented one of the methods by which this is accomplished. They can then be made heavier and with a larger surface of contact, so that the leakage is diminished, and the limit of the useful action of the pump is much extended. With the best pumps of this sort a pressure of one-half a millimetre of mercury is reached.

The *Sprengel air-pump* depends for its action upon the principle, discussed in § 93, that a stream of liquid running down a cylinder diminishes the pressure upon its walls. In the

Sprengel pump the liquid used is mercury. It runs from a large vessel down a glass tube, into the wall of which, at a distance from the bottom of the tube of more than 760 millimetres, enters the tube which connects with the receiver. The lower end of the vertical tube dips into mercury, which prevents air from passing up along the walls of the tube. When the stream of mercury first begins to flow, the air enters the column from the receiver, in consequence of the diminished pressure, passes down with the mercury in large bubbles, and emerges at the bottom of the tube. As the exhaustion proceeds, the bubbles become smaller and less frequent, and the mercury falls in the tube with a sharp, metallic sound. It is evident that, as in the case of the ordinary air-pump, a perfect vacuum cannot be secured. There is no leakage, however, in this form of the air-pump, and a very high degree of exhaustion can be reached.

The Morren or Alvergniat *mercury-pump* is in principle merely a common air-pump, in which combinations of stop-cocks are used instead of valves, and a column of mercury in place of the piston. Its particular excellence is that there is scarcely any leakage.

The *compressing-pump* is used, as its name implies, to increase the density of air or any other gas within a receiver. The receiver in this case is generally a strong metallic vessel. The working parts of the pump are precisely those of the air-pump, with the exception that the valves open downwards. As the piston is raised, air enters the cylinder, and is forced into the receiver at the down-stroke.

96. Manometers.—The *manometer* is an instrument used for measuring pressures. One variety depends for its operation upon the regularity of change of volume of a gas with change of pressure. This, in its typical form, consists of a heavy glass tube of uniform bore, sealed at one end, with the open end immersed in a basin of mercury. The pressure to be measured is applied to the surface of the mercury in the basin. As this pressure increases, the air contained in the tube is compressed, and a column of mercury is forced up the tube. The top of this column serves as an index. We know, from Boyle's law, that, when the volume of the air has diminished one-half, the pressure is doubled. The downward pressure of the mercury column makes up a part of this pressure; and the pressure acting on the surface of the mercury in the basin is greater than that indicated by the compression of the air in the tube, by the pressure due to the mercury column. For many purposes the manometer tube may be made very short, and the pressure of the mercury column that rises in it may be neglected.

97. **Aneroids.**—The *aneroid* is an instrument used to determine ordinary atmospheric pressures. On account both of its delicacy and its easy transportability, it is often used instead of the barometer. It consists of a metallic box, the cover of which is made of thin sheet-metal corrugated in circular grooves. The air is partially exhausted from the box, and it is then sealed. Any change in the pressure of the atmosphere causes the corrugated top to move. This motion is very slight, but is made perceptible, either by a combination of levers, which amplifies it, or by an arm rigidly fixed on the top, the motion of which is observed by a microscope. The indications of the aneroid are compared with those of a standard mercurial barometer, and an empirical scale is thus made, by means of which the aneroid may be used to determine pressures directly.

98. **Limitations to the Accuracy of Boyle's Law.**—In all the previous discussions, we have dealt with gases as if they obeyed Boyle's law with absolute exactness. This, however, is not the case. In the first place, some gases at ordinary temperatures can be liquefied by pressure. As these gases approach more nearly the point of liquefaction, the product pv

of the volume and pressure becomes less than it ought to be in accordance with Boyle's law.

Secondly, those gases which cannot be liquefied at ordinary temperatures by any pressure, however great, show a different departure from the law. For every gas, except hydrogen, there is a minimum value of the product pv. At ordinary temperatures and small pressures the gas follows Boyle's law quite closely, becoming, however, more compressible as the pressure increases, until the minimum value of pv is reached. It then becomes gradually less compressible, and at high pressures its volume is much greater than that determined by Boyle's law. If the temperature be raised, the agreement with the law is closer, and the pressure at which the minimum value of pv occurs is greater. Hydrogen seems to differ from the other gases, only in that the pressures at which the observations upon it were made were probably greater than the one at which its minimum value of pv occurs. The volume of the compressed hydrogen is uniformly greater than that required by Boyle's law.

Important modifications are introduced into the behavior of gases under pressure by subjecting them to intense cold. It is then found that all gases, without exception, can be liquefied, and even solidified.

The subject is intimately connected with the subject of critical temperature, and will be again discussed under Heat.

HEAT.

CHAPTER I.

MEASUREMENT OF HEAT.

99. **General Effects of Heat.**—Bodies are warmed, or their temperature is raised, by heat. The sense of touch is often sufficient to show difference in temperature; but the true criterion is the transfer of heat from the hotter to the colder body when the two bodies are brought in contact, and no work is done by one upon the other. This transfer is known by some of the effects described below.

Bodies, in general, expand when heated. Experiment shows that different substances expand differently for the same rise of temperature. Gases, in general, expand more than liquids, and liquids more than solids. Expansion, however, does not universally accompany rise of temperature. A few substances contract when heated.

Heat changes the state of aggregation of bodies, always in such a way as to admit of greater freedom of motion among the molecules. The melting of ice and the conversion of water into steam are familiar examples.

Heat breaks up chemical compounds. The compounds of sodium, potassium, lithium, and other metals, give to the flame of a Bunsen lamp the characteristic colors of the vapors of the metals which they contain. This fact shows that the heat separates the metals from their combinations.

When the junction of two dissimilar metals in a conducting circuit is heated, electric currents are produced.

Heat performs mechanical work. For example, the heat produced in the furnace of a steam-boiler may be used to drive an engine.

100. **Production of Heat.**—Heat is produced by various processes, some of which are the reverse of the operations just mentioned as the effects of heat. As examples of such reverse operations may be mentioned, the production of heat by the compression of a body which expands when heated; the production of heat during a change in the state of aggregation of a body, when the freedom of motion among the molecules is diminished; the production of heat during chemical combination; and the production of heat when an electric current passes through a junction of two dissimilar metals in an opposite direction to that of the current which is set up when the junction is heated.

Heat is produced in general in any process involving the expenditure of mechanical energy. The heat produced in such processes cannot be used to restore the whole of the original mechanical energy. The production of heat by friction is the best example of these processes.

Further, an electric current, in a homogeneous conductor, generates heat at every point in it, while, if every point in the conductor be equally heated, no current will be set up.

These cases are examples of the production of heat by non-reversible processes.

101. **Nature of Heat.**—Heat was formerly considered to be a substance which passed from one body to another, lowering the temperature of the one and raising that of the other, which combined with solids to form liquids, and with liquids to form gases or vapors. But the most delicate balances fail to show any change of weight when heat passes from one body to another. Rumford was able to raise a considerable quan-

tity of water to the boiling-point by the friction of a blunt boring-tool within the bore of a cannon. He showed that the heat manifested in this experiment could not have come from any of the bodies present, and also that heat would continue to be developed as long as the borer continued to revolve, or that the supply of heat was practically inexhaustible. The heat, therefore, must have been generated by the friction.

That ice is not melted by the combination with it of a heat substance was shown early in the present century by Davy. He caused ice to melt by friction of one piece upon another in a vacuum, the experiment being performed in a room where the temperature was below the melting-point of ice. There was no source from which heat could be drawn. The ice must, therefore, have been melted by the friction.

Rumford was convinced that the heat obtained in his experiment was only transformed mechanical energy; but to demonstrate this it was necessary to prove that the quantity of heat produced was always proportional to the quantity of mechanical work done. This was done in the most complete manner by Joule in a series of experiments extending from 1842 to 1849. He showed, that, however the heat was produced by mechanical means, whether by the agitation of water by a paddle-wheel, the agitation of mercury, or the friction of iron plates upon each other, the same expenditure of mechanical energy always developed the same quantity of heat. Joule also proved the perfect equivalence of heat and electrical energy.

These experiments prove that *heat is a form of energy.* Consistent explanations of all the phenomena of heat may be given if we assume that the molecules of all bodies are in constant motion, that the temperature of a body varies with the mean kinetic energy of its molecules, and that the heat in a body is the sum of the kinetic energies of its molecules.

THERMOMETRY.

102. Temperature.—Two bodies are said to be at the same *temperature* when, if they be brought into intimate contact, no heat is transferred from one to the other. A body is at a high temperature relatively to other bodies when it gives up heat to them. The fact that it gives up heat may be shown by its change in volume. A body is at a low temperature when it receives heat from surrounding bodies. It is understood, of course, in what is said above, that one body has no action upon the other; in other words, no work is done by one body upon the other when they are brought in contact.

103. Thermometers.—Experiments show that, in general, bodies expand, and their temperature rises progressively, with the application of heat. An instrument may be constructed which will show at any instant the volume of a body selected for the purpose. If the volume increase, we know that the temperature rises; if the volume remain constant or diminish, we know that the temperature remains stationary or falls. Such an instrument is called a *thermometer*.

The thermometer most in use consists of a glass bulb with a fine tube attached. The bulb and part of the tube contain mercury. In order that the thermometers of different makers may give similar readings, it is necessary to adopt two standard temperatures which can be easily and certainly reproduced. The temperatures adopted are the melting-point of ice, and the temperature of steam from boiling water, under a pressure equal to that of a column of mercury 760 millimetres high at Paris. After the instrument has been filled with mercury, it is plunged in melting ice, from which the water is allowed to drain away, and a mark is made upon the stem opposite the end of the mercury column. It is then placed in a vessel in which water is boiled, so constructed that the steam rises through a tube surrounding the thermometer, and then descends by an annular space between that tube and an outer one, and escapes at the bottom. The thermometer does not touch the water, but is entirely surrounded by steam. The point reached by the end of the mercury column is marked on the stem, as before. The space between these two marks is then divided into a number of equal parts.

While all makers of thermometers have adopted the same standard temperatures for the fixed points of the scale, they differ as to the number of divisions between these points. The thermometers used for scientific purposes, and in general use in France, have the space between the fixed points divided into a hundred equal parts or *degrees*. The melting-point of ice is marked 0°, and the boiling-point 100°. This scale is called the *Centigrade* or *Celsius* scale.

The *Réaumur* scale, in use in Germany, has eighty degrees between the melting- and boiling-points, and the boiling-point is marked 80°.

The *Fahrenheit* scale, in general use in England and America, has a hundred and eighty degrees between the melting- and boiling-points. The former is marked 32°, and the latter 212°.

The divisions in all these cases are extended below the zero point, and are numbered from zero downward. Temperatures below zero must, therefore, be read and treated as negative quantities.

A few points in the process of construction of a thermometer deserve notice. It is found that glass, after it has been heated to a high temperature and again cooled, does not for some time return to its original volume. The bulb of a thermometer must be heated in the process of filling with mercury, and it will not return to its normal volume for some months. The construction of the scale should not be proceeded with until the reservoir has ceased to contract. For the same reason, if the thermometer be used for high temperatures, even the temperature of boiling water, time must be given for the reservoir to return to its original volume before it is used for the measurement of low temperatures.

It is essential that the diameter of the tube should be nearly uniform throughout, and that the divisions of the scale should represent equal capacities in the tube. To test the tube a thread of mercury about 50 millimetres long is introduced, and its length is measured in different parts of the tube. If the length vary by more than a millimetre, the tube should be rejected. If the tube be found to be suitable, a bulb is attached, mercury is introduced, and the tube sealed after the mercury has been heated to expel the air. When it is ready for graduation, the fixed points are determined; then a thread of mercury having a length equal to about ten degrees of the scale is detached from the column, and its length measured in all parts of the tube. By reference to these measurements, the tube is so graduated that the divisions represent parts of equal capacity, and are not necessarily of equal length.

If such a thermometer indicate a temperature of 10°, this means that the thermometer is in such a thermal condition that the volume of the mercury has increased from zero one tenth of its total expansion from zero to 100°. There is no reason for supposing that this represents the same proportional rise of temperature. If a thermometer be constructed in the manner described, using some liquid other than mercury, it will not in general indicate the same temperature as the mercurial thermometer, except at the two standard points. It is plain, therefore, that a given fraction of the expansion of a liquid from zero to 100° cannot be taken as representing the same fraction of the rise of temperature.

104. Air Thermometer.—If a gas be heated, and its volume kept constant, its pressure increases. For all the so-called permanent gases—that is, those which are liquefied only with great difficulty—the amount of increase in pressure for the same increase of temperature is found to be almost exactly the same. This fact is a reason for supposing that the increase of pressure is proportional to the increase of temperature. There are theoretical reasons, as will be seen later, for the same supposition.

An instrument constructed to take advantage of this increase in pressure to measure temperature is called an *air thermometer.* A bulb so arranged that it may be placed in the medium of which the temperature is to be determined, is filled with air or some other gas, and means are provided for maintaining the volume of the gas constant, and measuring its pressure. For the reasons given above, the air thermometer is taken as the standard instrument for scientific purposes. Its use, however, involves several careful observations and tedious computations. It is, therefore, mainly employed as a standard with which to compare other instruments. If we make such a comparison, and construct a table of corrections, we may reduce the readings of any thermometer to the corresponding readings of the air thermometer.

105. Limits in the Range of the Mercurial Thermometer.—The range of temperature for which the mercurial thermometer may be employed is limited by the freezing of the mercury on the one hand, and its boiling on the other. For temperatures below the freezing-point of mercury, alcohol thermometers may be employed. For the measurement of high temperatures, several different methods have been employed. One depends upon the expansion of a bar of platinum, another upon the variation in the electric resistance of platinum wire, another upon the strength of the electric current generated by a thermo-electric pair, another on the density of mercury vapor. This last method is carried out as follows: A globe of refractory material, fitted with a short tube with a small bore, contains a small quantity of mercury. It is placed in the furnace or other place, the temperature of which is to be measured. The mercury boils, the air and the excess of mercury are expelled, and the globe is finally left full of mercury vapor at the temperature of the furnace. The globe is cooled, and the weight of the mercury left in it determined. From this and the volume of the globe the temperature can be computed.

106. Registering Thermometers.—*Maximum* and *minimum thermometers* are employed to register the highest and lowest temperatures reached during a given period. By a change in construction, the ordinary mercury thermometer becomes a self-registering maximum thermometer. This change consists in making a contraction in the tube just above the reservoir, to such an extent that, though the mercury is pushed through it as the temperature rises, it does not return as the temperature falls. It thus serves as an index to show the highest temperature reached during the period of its exposure. After an observation has been made, the thermometer is readjusted for a new observation by allowing the instrument to swing out of the horizontal position in which it usually rests, about a point near the upper extremity of the tube.

In the construction of the minimum thermometer, alcohol is the liquid employed. Before sealing, an index of glass, smaller in diameter than the bore of the tube, is inserted. When the instrument is adjusted for use, this is brought in contact with the extremity of the column, and the tube is placed in a horizontal position. If, now, the alcohol expand, it will flow past the index without moving it; but if it contract, it will, by adhesion, draw the index after it. The minimum temperature is thus registered.

Registering thermometers have been made to give a continuous record of changes of temperature. One method of effecting this is to produce an image of the thermometer tube, which is strongly illuminated by a light placed behind it, upon a screen of sensitized paper which moves continuously by means of clockwork. Light is excluded from the whole of the paper, except the part that corresponds to the image of the tube above the mercury. This part of the paper is blackened by the light; and, as the paper moves, the edge of the blackened portion will present a sinuous line corresponding to the movements of the mercury of the thermometer.

CALORIMETRY.

107. Unit of Heat.—It is evident that more heat is required to raise the temperature of a large quantity of a substance through a given number of degrees than to raise the temperature of a small quantity of the same substance through the same number of degrees. It is further evident that the successive repetition of any operation by which heat is produced will generate more heat than a single operation. Heat is therefore a quantity the magnitude of which may be expressed in terms of some unit. The unit of heat generally adopted is the heat required to raise the temperature of one kilogram of water from zero to one degree. It is called a *calorie.*

It is sometimes convenient to employ a smaller unit, namely, the quantity of heat necessary to raise one gram of water from zero to one degree. This unit is designated as the *lesser calorie.* It is one one-thousandth of the larger unit. It may, therefore, be called a *millicalorie.*

The fact that heat is energy enables us to employ still another unit. It is that quantity of heat which is equivalent to an erg. This unit is called the *mechanical unit* of heat. A calorie contains about 41,595,000,000 mechanical units.

108. Heat required to raise the Temperature of a Mass of Water.—It is evident that to raise the temperature of m kilograms of water from zero to one degree will require m calories. If the temperature of the same quantity of water fall from one degree to zero, the same quantity of heat is given to surrounding bodies.

Experiment shows, that, if the same quantity of water be raised to different temperatures, quantities of heat nearly proportional to the rise in temperature will be required: hence, to raise the temperature of m kilograms of water from zero to t degrees requires mt calories very nearly. This is shown by mixing water at a lower temperature with water at a higher temperature. The temperature of the mixture will be almost exactly the mean of the two. Regnault, who tried this experiment with the greatest care, found the temperature of the mixture a little higher than the mean, and concluded that the quantity of heat required to raise the temperature of a kilogram of water one degree increases slightly with the temperature; that is, to raise the temperature of a kilogram of water from twenty to twenty-one degrees, requires a little more heat than to raise the temperature of the same quantity of water from zero to one degree.

Rowland found, by mixing water at various temperatures, and also by measuring the energy required to raise the temperature of water by agitation by a paddle-wheel, that, when the air thermometer is taken as a standard, the quantity of heat necessary to raise the temperature of a given quantity of water one degree diminishes slightly from zero to thirty degrees, and then increases to the boiling-point.

109. Specific Heat.—Only one thirtieth as much heat is required to raise the temperature of a kilogram of mercury from zero to one degree as is required to raise the temperature of a kilogram of water through the same range. In order to raise the temperatures of other substances through the same range, quantities of heat peculiar to each substance are required.

The quantity of heat required to raise the temperature of one kilogram of a substance from zero to one degree is called the *specific heat* of the substance.

If the temperature of one kilogram of a substance rise from t_1 to t, the limit of the ratio of the quantity of heat required to bring about the rise in temperature to the difference in temperature, as that difference diminishes indefinitely, is called the specific heat of the substance at temperature t. If we represent the quantity of heat by Q, the limit of the ratio $\frac{Q}{t - t_{\prime}} = \frac{dQ}{dt}$ expresses this specific heat.

The specific heats of substances are generally nearly constant between zero and one hundred degrees. The *mean specific heat* of a substance between zero and one hundred degrees is the one usually given in the tables.

The measurement of specific heat is one of the important objects of *calorimetry*.

110. Ice Calorimeter.—*Black's ice calorimeter* consists of a block of pure ice having a cavity in its interior covered by a thick slab of ice. The body of which the specific heat is to be determined is heated to t degrees, then dropped into the cavity, and immediately covered by the slab. After a short time the temperature of the body falls to zero, and in so doing converts a certain quantity of ice into water. This water is removed by a sponge of known weight, and its weight is determined. It will be shown, that to melt a kilogram of ice requires 80 calories; if, then, the weight of the body be P, and its specific heat x, it gives up, in falling from t degrees to zero, Pxt calories. On the other hand, if p kilograms of ice be melted, the heat required is $80p$. Therefore $Pxt = 80p$; whence

$$x = \frac{80p}{Pt}. \tag{38}$$

Bunsen's ice calorimeter (Fig. 44) is used for determining the specific heats of substances of which only a small quantity is at hand. The apparatus is entirely of glass. The tube B is filled with water and mercury, the latter extending into the graduated capillary tube C. To use the apparatus, alcohol which has been artificially cooled to a temperature below zero is passed through the tube A. A layer of ice forms around the outside of this tube. As water freezes, it expands. This causes the mercury to advance in the capillary tube C. When a sufficient quantity of ice has been formed, the alcohol is removed from A, the apparatus is surrounded by melting snow or ice, and a small quantity of water is introduced, which soon falls in temperature to zero. The position of the mercury in C is now noted; and the substance the specific heat of which is to be determined, at the temperature of the surrounding air, is dropped into the water in A.

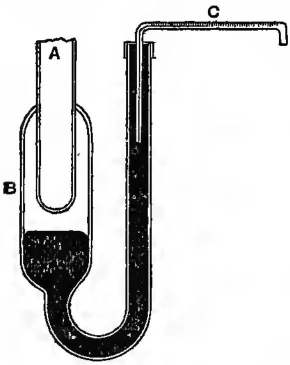

FIG. 44.

Its temperature quickly falls to zero, and the heat which it loses is entirely employed in melting the ice which surrounds the tube A. As the ice melts, the mercury in the tube C retreats. The change of position is an indication of the quantity of ice melted, and the quantity of ice melted measures the heat given up by the substance. The number of divisions of the tube C corresponding to one calorie can be determined by direct experiment. To make this determination, the operation is performed as described above, using a substance of a known specific heat c. If its mass be p and its temperature t, it gives up, in cooling to zero, cpt calories. If the mercury retreat at the same time n divisions, one division corresponds to $\frac{cpt}{n}$ calories. If, now, a mass p' of a substance at a temperaature t' be introduced, and the mercury fall n' divisions, the number of calories which must have been given to the ice is $\frac{n'cpt}{n}$, and the specific heat of the substance is

$$x = \frac{n'cpt}{np't'}. \tag{39}$$

111. Method of Mixtures.—The method of mixtures consists in bringing together, at different temperatures, the substance of which the specific heat is desired and another of which the specific heat is known, and noting the change of temperature which each undergoes.

The *water calorimeter* consists of a vessel of very thin copper or brass, highly polished, and placed within another vessel upon non-conducting supports. A mass P of the substance of which the specific heat is to be determined is brought to a temperature $t_{,}$ in a suitable bath, then plunged in water at the temperature t, contained in the calorimeter. The whole will soon come to a common temperature θ.

The substance loses $Px\,(t_{,} - \theta)$ calories.

The heat gained by the calorimeter consists of the heat gained by the water, $p(\theta - t)$; the heat gained by the vessel, $p'c'\,(\theta - t)$; the heat gained by the glass stirrer and the glass of the thermometer, $p''c''\,(\theta - t)$; the heat gained by the mercury of the thermometer $p'''c'''\,(\theta - t)$: where p represents the mass of the water, p' the mass of the vessel, p'' the mass of the glass of the stirrer and of the thermometer, p''' the mass of the mercury of the thermometer, c' the specific heat of the material of the vessel, c'' the specific heat of glass, and c''' the specific heat of mercury. If no heat be lost or gained by radiation, the heat lost by the substance is equal to that gained by the calorimeter:

whence $$Px(t_{,} - \theta) = (p + p'c' + p''c'' + p'''c''')(\theta - t). \quad (40)$$

To determine x from this formula, c', c'', and c''' must be known. Approximate values of these may be obtained and used in the formula, but it is better to determine the value of $p'c' + p''c'' + p'''c''' = p_{,}$ by experiment. Let a mass of water, P, at a temperature $t_{,}$ be substituted for the substance of which the specific heat was to be determined in the experiment described above. The equation will then become

$$P(t_{,} - \theta) = (p + p_{,})(\theta - t), \quad (41)$$

in which $p_{\prime}$ is the only unknown quantity. $p_{\prime}$ is the *water equivalent* of the calorimeter and accessories. It is determined, once for all, as just described.

There is a source of error in the use of the instrument, due to the radiation of heat during the experiment. This error may be nearly eliminated, by making a preliminary experiment to determine what change of temperature the calorimeter will experience; then, for the final experiment, the calorimeter and its contents are brought to a temperature below the temperature of the surrounding air, by about half the amount of that change. The calorimeter will then receive heat from the surrounding medium during the first part of the experiment, and lose heat during the second part. The rise of temperature is, however, much more rapid at the beginning than at the end of the experiment. The rise from the initial temperature to the temperature of the surrounding medium occupies less time than the rise from the latter to the final temperature. The gain of heat, therefore, does not exactly compensate for the loss. If greater accuracy be required, the rate of cooling of the calorimeter must be determined by putting into it warm water, the same in quantity as would be used in experiments for determining specific heat, and noting its temperature from minute to minute. Such an experiment furnishes the data for computing the loss or gain by radiation. To secure accurate results the body must be transferred from the bath to the calorimeter without sensible loss of heat.

112. **Method of Comparison.**—The method of comparison consists in conveying to the substance of which the specific heat is to be determined a known quantity of heat, and comparing the consequent rise of temperature with that produced by the same amount of heat in a substance of which the specific heat is known. In the early attempts to use this method, the heat produced by the same flame burning for a given time was applied successively to different liquids. A more exact

method was the combustion, within the calorimeter, of a known weight of hydrogen. The best method of obtaining a known quantity of heat is by means of an electrical current of known strength flowing through a wire of known resistance wrapped upon the calorimeter.

113. Method of Cooling.—The method of cooling consists in noting the time required for the calorimeter, in a space kept constantly at zero, to cool from a temperature t' to a temperature t, when empty; when containing a given weight of water; and when containing a given weight of the substance of which the specific heat is sought. The *thermo-calorimeter* of Regnault, represented in Fig. 45, is an example. It consists of an alcohol thermometer, with its bulb A enlarged and made in the form of a hollow cylinder, inside of which the substance is placed. The thermometer is warmed, and then placed in a vessel surrounded by melting ice. It radiates heat to the sides of the vessel, and the column of alcohol in the tube falls. Let x be the time occupied in falling from the division n to the division n' when the space B is empty. Let the times occupied in falling between the same two divisions, when the space B contains a mass P of water, and when it contains a mass P' of the substance of which the specific heat c' is sought, be respectively x' and x''. Let M be the water equivalent of the instrument. We then have

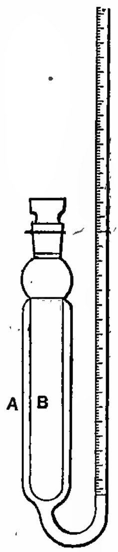

FIG. 45.

$$\frac{M}{x} = \frac{M+P}{x'} = \frac{M+P'c'}{x''},$$

since, under the conditions of the experiment, the heat lost per second must be the same in each case.

Eliminating M, we obtain

$$c' = \frac{P}{P'}\left(\frac{x'' - x}{x' - x}\right). \qquad (42)$$

114. Determination of the Mechanical Equivalent of Heat.—It has been stated that whenever heat is produced by the expenditure of mechanical energy, the quantity of heat produced is always proportional to the quantity of mechanical energy expended.

The *mechanical equivalent* of heat is the energy in mechanical units, the expenditure of which produces the unit of heat.

Heat applied to a body may increase the motion of its molecules; that is, add to their kinetic energy. It may perform internal work by moving the molecules against molecular forces. It may perform external work by producing motion against external forces. If we could estimate these effects in mechanical units, we might obtain the mechanical equivalent of heat. But the kinetic energy of the molecules cannot be estimated, for we do not know their mass nor their velocity. We must, therefore, in the present state of our knowledge, resort to direct experiment to determine the heat equivalent. In one of the experiments of Joule, already referred to, a paddle-wheel was made to revolve, by means of weights, in a vessel filled with water. In this vessel were stationary wings, to prevent the water from acquiring a rotary motion with the paddle-wheel. By the revolution of the wheel the water was warmed. The heat so generated was estimated from the rise of temperature, while the mechanical energy required to produce it was given by the fall of the driving-weight. Joule repeated this experiment, substituting mercury for the water. In another experiment he substituted an iron plate for the paddle-wheel, and made it revolve with friction upon a fixed iron plate under water.

Joule expressed his results in kilogram-metres—that is, the work done by a kilogram in falling under the force of gravity through one metre. He stated the mechanical equivalent of one calorie, in this unit, to be 423.9, from the experiments with water; 425.7, from those with mercury; and 426.1, from those with iron plates. He gave the preference to the

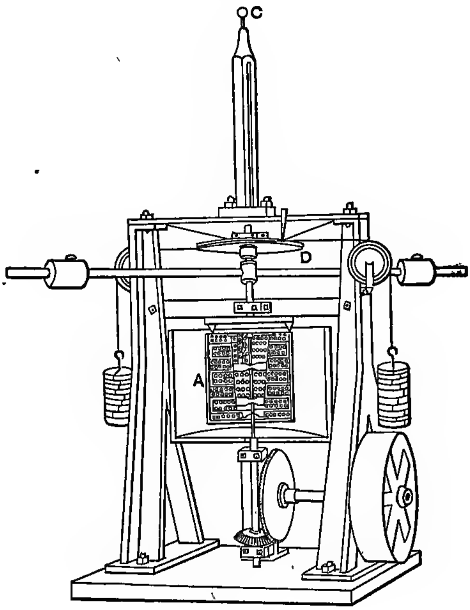

FIG. 46.

smallest value, and it has been generally accepted as the mechanical equivalent. This mechanical equivalent is called *Joule's equivalent*, and is represented by J. In absolute units it is about 41,595,000,000 ergs per calorie.

Rowland has repeated Joule's experiment with water; but he caused the paddle-wheel to revolve by means of an engine, and determined the moment of the couple required to prevent the revolution of the calorimeter. Fig. 46 shows the apparatus. The shaft of the paddle-wheel projects through the bottom of the calorimeter, and is driven by means of a bevel-gear. The vessel *A* is suspended from *C* by a torsion wire, and its tendency to rotate balanced by weights attached to cords, which act upon the circumference of a pulley *D*. By this disposition of the apparatus he was able to expend about one half a horse-power in the calorimeter, and obtain a rise of temperature of 35° per hour; while in Joule's experiments the rise of temperature per hour was less than 1°. These experiments give, for the mechanical equivalent of one calorie at 5°, 429.8 kilogram-metres; at 20°, 426.4 kilogram-metres.

Several other methods have been employed for determining the mechanical equivalent. The concordance of the results by all these methods is sufficient to warrant the statement that the expenditure of a given amount of mechanical energy always produces the same amount of heat.

CHAPTER II.

TRANSFER OF HEAT.

115. **Transfer of Heat.**—In the preceding discussions it has been assumed that heat may be transferred from one body to another, and that if two bodies in contact be at different temperatures, heat will be transferred from the hotter to the colder body. In general, if transfer of heat be possible in any system, heat will pass from the hotter to the colder parts of the system, and the temperature of the system will tend to become uniform. There are three ways in which this transfer is accomplished, called respectively convection, conduction, and radiation.

116. **Convection.**—If a vessel containing any fluid be heated at the bottom, the bottom layers become less dense than those above, producing a condition of instability. The lighter portions of the fluid rise, and the heavier portions from above, coming to the bottom, are in their turn heated. Hence continuous currents are caused. This process is called *convection.* By this process, masses of fluid, although fluids are poor conductors, may be rapidly heated. Water is often heated in a reservoir at a distance from the source of heat by the circulation produced in pipes leading to the source of heat and back. The winds and the great currents of the ocean are convection currents. An interesting result follows from the fact that water has a maximum density (§ 135). When the water of lakes cools in winter, currents are set up and maintained, so long as the surface water becomes more dense by cooling, or until the whole mass reaches 4°. Any further cooling makes the

surface water lighter. It therefore remains at the surface, and its temperature rapidly falls to the freezing-point, while the great mass of the water remains at the temperature of its maximum density.

117. Conduction.—If one end of a metal rod be heated, it is found that the heat travels along the rod, since those portions at a distance from the source of heat finally become warm. This process of transfer of heat from molecule to molecule of a body, while the molecules themselves retain their relative places, is called *conduction.*

In the discussion of the transfer of heat by conduction it is assumed as a principle, borne out by experiment, that the flow of heat between two very near parallel planes, drawn in a substance, is proportional to the difference of temperature between those planes.

118. Flow of Heat across a Wall.—To study the transfer of heat by conduction, we will consider what takes place in a wall of homogeneous material, the exposed surfaces of which, assumed to be of indefinite extent, are maintained at a constant difference of temperature. Suppose the wall to be cut by a series of planes parallel to the exposed surfaces; and that the state of the body, as respects temperature, has become permanent. Then we will show that there must be the same flow of heat across all parallel sections, and also that there must be a uniform fall of temperature from one side of the wall to the other,—that is, that if $t'-t$ represent the difference of temperature between the two exposed surfaces, and d the thickness of the wall, $\frac{t'-t}{d}$ is the fall of temperature per unit thickness, and $t' - \frac{t'-t}{d}d'$ is the temperature at a distance d' from the warmer surface.

To demonstrate this, suppose A (Fig. 47) to be one exposed surface at temperature t', and B the other surface at temperature t: suppose a, a', a'', to be three surfaces parallel to the faces of the wall, and at very small equal distances from one another. Suppose the temperatures to exist according to the law stated in the proposition: then the difference of temperature between a and a' will be the same as between a' and a''. It has been stated that the flow of heat between two points in a body is proportional to the difference of temperature between those points. Experiment shows that it depends also upon the distance between them, the nature of the material, and to a very limited extent upon the temperature itself. The effect of this last factor may, however, be neglected, since the pairs of surface considered are nearly at the same temperature. The other factors being the same for both pairs, it follows that there will be the same flow of heat from a to a' as from a' to a''. The same will hold true for any other set of surfaces parallel to the faces of the wall: hence the molecules in any surface such as a' receive and part with equal amounts of heat, and can neither rise nor fall in temperature. If the temperatures, therefore, were once established in accordance with the law enunciated, they could never change. On the other hand, if the difference of temperature between a and a' were greater than that between a' and a'', the molecules in a' would receive more heat than they would part with, their temperature would rise, and this would tend to equalize these differences. The proposition is therefore demonstrated.

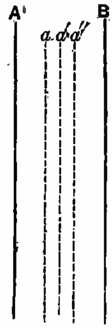

FIG. 47.

119. Flow Proportional to Rate of Fall of Temperature.—It can be further shown that the flow of heat across walls of the same material is directly proportional to the differences of temperature between the faces of the walls, and inversely proportional to the thicknesses of the walls. Represent the thicknesses of two walls A and B by d and δ respectively, the temperatures of the exposed surfaces of A by t' and t, and of B by θ' and θ. Assume two planes in each wall parallel to the exposed surfaces, at very small equal distances apart, and similarly situated. The flow of heat in each wall, from the one plane to the other, will be proportional only to the differences of temperature, since all other things are equal. If ϵ be the common distance between the planes, $\frac{\epsilon}{d}(t' - t)$ will be the difference of temperature between the planes in A, and $\frac{\epsilon}{\delta}(\theta' - \theta)$ will be the same in B; hence we have

$$\frac{\textit{flow of heat between the planes in } A}{\textit{flow of heat between the planes in } B} = \frac{\frac{\epsilon}{d}(t' - t)}{\frac{\epsilon}{\delta}(\theta' - \theta)} = \frac{\frac{t' - t}{d}}{\frac{\theta' - \theta}{\delta}}.$$

This proves the proposition, since the heat which flows across the wall is the same as that which flows between any two planes.

It will be seen that $\frac{t' - t}{d}$ is the *rate of fall of temperature* at the section considered; and it follows, finally, that the flow of heat across any section parallel to the exposed surfaces of a wall is proportional to the rate of fall of temperature at that section.

120. **Conductivity.**—If, now, we consider a prism extending across the wall, bounded by planes perpendicular to the exposed surfaces, and represent the area of its exposed bases by A, the quantity of heat which flows in a time T through this prism may be represented by

$$Q = K\frac{t' - t}{d}AT, \tag{43}$$

where K is a constant depending upon the material of which the wall is composed. K is the *conductivity* of the substance, and may be defined as the quantity of heat which in unit time flows through a section of unit area in a wall of the substance whose thickness is unity, when its exposed surfaces are maintained at a difference of temperature of one degree; or, in other words, it is the quantity of heat which in unit time flows through a section of unit area in a substance, where the rate of fall of temperature at that section is unity. In the above discussions the temperatures t' and t are taken as the actual temperatures of the surfaces of the wall. If the colder surface of the wall be exposed to air of temperature T, to which the heat which traverses it is given up, t will be greater than T. The difference will depend upon the quantity of heat which flows, and upon the facility with which the surface parts with heat.

121. **Flow of Heat along a Bar.**—If a prism of a substance have one of its bases maintained at a temperature t, while the other base and the sides are exposed to air at a lower temperature, the conditions of uniform fall of temperature no longer exist, and the amount of heat which flows through the different sections is no longer the same; but the amount of heat which flows through any section is still proportional to the rate of fall of temperature at that section, and is equal to the heat which escapes from the portion of the bar beyond the section.

122. **Measurement of Conductivity.**—A bar heated at one end furnishes a convenient means of measuring conductivity. In Fig. 48 let AB represent a bar heated at A. Let the ordinates aa', bb', cc', represent the excess of temperatures above the temperature of the air at the points from which they are drawn. These temperatures may be determined by means of

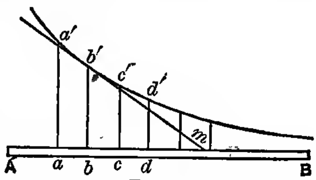

Fig. 48.

thermometers inserted in cavities in the bar, or by means of a thermopile. Draw the curve $a'b'c'd' \ldots$ through the summits of the ordinates. The inclination of this curve at any point represents the rate of fall of temperature at that point. The ordinates to the line $b'm$, drawn tangent to the curve at the point b', show what would be the temperatures at various points of the bar if the fall were uniform and at the same rate as at b'. It shows that, at the rate of fall at b', the bar would at m be at the temperature of the air; or, in the length bm, the fall of temperature would equal the amount represented by bb'. The rate of fall is, therefore, $\frac{bb'}{bm}$. If Q represent the quantity of heat passing the section at b in the unit time, we have, from § 120,

$$Q = K \times \textit{rate of fall of temperature} \times \textit{area of section}.$$

Q is equal to the quantity of heat that escapes in unit time from all that portion of the bar beyond b. It may be found by heating a short piece of the same bar to a high temperature, allowing it to cool under the same conditions that surround the bar AB, and observing its temperature from minute to minute as it falls. These observations furnish the data for computing the quantity of heat which escapes per minute from unit length of the bar at different temperatures. It is then easy to compute the amount of heat that escapes per minute from each portion, bc, cd, etc., of the bar beyond b; each portion being taken so short that its temperature throughout may, without sensible error, be considered uniform and the same as that at its middle point. Summing up all these quantities, we obtain the quantity Q which passes the section b in the unit time. Then

$$K = \frac{Q}{\textit{rate of fall of temperature at } b \times \textit{area of section}}.$$

123. Conductivity diminishes as Temperature rises.—By the method described above, Forbes determined the conductivity of a bar of iron at points at different distances from the heated end, and found that the conductivity is not the same at all temperatures, but is greater as the temperature is lower.

124. Conductivity of Crystals.—The conductivity of crystals of the isometric system is the same in all directions, but in crystals of the other systems it is not so. In a crystal of Iceland spar the conductivity is greatest in the direction of the axis of symmetry, and equal in all directions in a plane at right angles to that axis.

125. Conductivity of Non-homogeneous Solids.—De la Rive and De Candolle were the first to show that wood conducts heat better in the direction of the fibres than at right angles to them. Tyndall, by experimenting upon cubes cut from wood, has shown that the conductivity has a maximum value parallel to the fibres, a minimum value at right angles to the fibres and parallel to the annual layers, and a medium value at right angles to both fibres and annual layers. Feathers, fur, and the materials of clothing, are poor conductors because of their want of continuity.

126. Conductivity of Liquids.—The conductivity of liquids can be measured, in the same way as that of solids, by noting the fall of temperature at various distances from the source of heat in a column of liquid heated at the top. Great care must be taken in these experiments to avoid errors due to convection currents.

Liquids are generally poor conductors.

127. Radiation.—We have now considered those cases in which there is a transfer of heat between bodies in contact. Heat is also transferred between bodies not in contact. This is effected by a process called *radiation*, which will be subsequently considered.

CHAPTER III.

EFFECTS OF HEAT.

SOLIDS AND LIQUIDS.

128. Expansion of Solids.—When heat is applied to a body it increases the kinetic energy of the molecules, and also increases the potential energy, by forcing the molecules farther apart against their mutual attractions and any external forces that may resist expansion. Since the internal work to be done when a solid or liquid expands varies greatly for different substances, it might be expected that the amount of expansion for a given rise of temperature would vary greatly.

In studying the expansion of solids, we distinguish *linear* and *voluminal* expansion.

The increase which occurs in the unit length of a substance for a rise of temperature from zero to 1° C. is called the *coefficient of linear expansion.* Experiment shows that the expansion for a rise of temperature of one degree is very nearly constant between zero and 100°.

Represent by l_0 the distance between two points in a body at zero, by l_t the distance between the same points at the temperature t, and by α the coefficient of linear expansion of the substance of which the body is composed.

The increase in the distance l_0 for a rise of one degree in temperature is αl_0, for a rise of t degrees $\alpha t l_0$. Hence we have, after a rise in temperature of t degrees,

$$l_t = l_0(1 + \alpha t), \qquad (44)$$

and $$l_0 = \frac{l_t}{1 + \alpha t};$$

or approximately, $$l_0 = l_t(1 - \alpha t).$$

The binomial $1 + \alpha t$ is called the *factor of expansion.*

In the same way, if k represent the *coefficient of voluminal expansion,* the volume of a body at a temperature t will be

$$V_t = V_0(1 + kt); \qquad (45)$$

and if d represent density, since density is inversely as volume, we have

$$d_t = \frac{d_0}{1 + kt}. \qquad (46)$$

For a homogeneous solid, the coefficient of voluminal expansion is three times that of linear expansion; for, if the temperature of a cube, with an edge of unit length, be raised one degree, the length of its edge becomes $1 + \alpha$, and its volume $1 + 3\alpha + 3\alpha^2 + \alpha^3$. Since α is very small, its square and cube may be neglected; and the volume of the cube after a rise in temperature of one degree is $1 + 3\alpha$. 3α is, therefore, the coefficient of voluminal expansion.

129. Measurement of Coefficients of Linear Expansion.—Coefficients of linear expansion are measured by comparing the lengths, at different temperatures, of a bar of the substance the coefficient of which is required, with the length, at constant temperature, of another bar. The constant temperature of the latter bar is secured by immersing it in melting ice. The bar, the coefficient of which is sought may be brought to different temperatures by immersing it in a liquid bath; but it is found better to place the bar upon the instrument by means of which the comparisons are to be made, and leave it for several hours exposed to the air of the room, which is kept at a constant temperature by artificial means. Of course several hours must elapse between any two comparisons by this method, and its application is restricted to such ranges of temperature as may be obtained in occupied rooms; but within this range the observations can be made much more accurately than would be the case when the bar is immersed in a bath, and it is within this range that an accurate knowledge of coefficients of expansion is of most importance.

130. Expansion of Liquids.—In studying the expansion of a liquid, it is important to distinguish its *absolute* expansion, or the real increase in volume, and its *apparent* expansion, or its increase in volume in comparison with that of the containing vessel.

131. Absolute Expansion of Mercury.—A knowledge of the coefficients of expansion of mercury is of the greatest importance, since mercury is made use of for so many purposes in physical research. Regnault has made the most accurate determinations of these constants.

To determine the absolute coefficient, the experiment must be so made that the expansion of the vessel shall not influence the result.

Regnault's method consists in comparing the heights of two columns of mercury, at different temperatures, which produce the same pressure. Two vertical tubes, ab, $a'b'$ (Fig. 49), are connected at the top by a horizontal tube aa', and at the bottom by a tube $bcdd'c'b'$, a part of which is of glass, and shaped like an inverted U. The top of the inverted U-tube is connected by a tube e with a vessel f, in which air can be maintained at any desired pressure. When these tubes are filled with mercury, it flows freely from one to the other at the top; but the flow of mercury between them at the bottom is prevented by air imprisoned in the U-tube, while the pressure is transmitted undiminished. The pressure at each end of the column of imprisoned air must, therefore, be the same; and, since a and a' are connected by a horizontal tube, the pressures at those points are the same also; hence the difference of pressure between a and d must be equal to the difference between a' and d'; and from § 86 it follows that the heights of the columns, without regard to the diameters of the tubes, producing this difference, are inversely proportional to their densities. If, now, one branch be raised to the temperature t, while the other remains at zero, the mercury in the U-tube will assume different levels. Measuring the height of each column from the surface of the mercury in the U-tube to the horizontal tube at the top, we have, from Eq. (46), if h and h' represent the height of the cold and warm columns respectively,

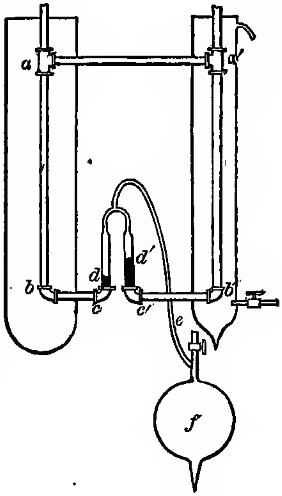

FIG. 49.

$$\left.\begin{aligned} \frac{h}{h'} &= \frac{d'}{d} = \frac{1}{1+kt}; \\ hkt &= h' - h; \\ k &= \frac{h'-h}{ht}. \end{aligned}\right\} \qquad (47)$$

132. Apparent Expansion of Mercury.—If a glass bulb (Fig. 50), furnished with a capillary tube, be filled with mercury at zero, and heated to a temperature t, some of the mercury runs out. The amount which overflows evidently depends upon the difference of expansion between the mercury and the glass. Let P represent the mass of mercury that fills the bulb and tube at zero. After heating, there remains in the bulb a mass $P-p$ of mercury, which at zero occupies the volume ab. The mass of mercury p, which runs out, would at the same temperature fill the remainder of the bulb and tube. Hence the volume ab equals $\frac{P-p}{d}$, where d represents the density of mercury. The volume above b equals $\frac{p}{d}$.

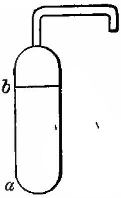

FIG. 50.

The mercury in ab, when heated to the temperature t, just fills the tube; and its apparent volume is $\frac{P}{d}$. If κ represent the coefficient of apparent expansion,

$$\frac{P}{d}=\frac{P-p}{d}(1+\kappa t), \tag{48}$$

or

$$\kappa=\frac{p}{(P-p)t}.$$

If we know κ, the instrument may be used as a thermometer; for, suppose it filled at zero, and subjected to an unknown temperature t_1, we shall have, it p_1 represent the mercury that then runs over,

$$P=(P-p_1)(1+\kappa t), \tag{49}$$

whence

$$t_1=\frac{p_1}{(P-p_1)\kappa}.$$

The instrument is, therefore, called a *weight thermometer.*

The difference between the value κ, found above, and the absolute coefficient k is due to the expansion of the glass. And if k be the coefficient for glass, we have

$$k' = k - \kappa;$$

for, referring again to Fig. 50, the volume of the vessel at zero is $\frac{P}{d}$, and at the temperature t is $\frac{P}{d}(1 + k't)$, which, from Eq. (48), equals

$$\frac{P-p}{d}(1 + \kappa t)(1 + k't).$$

The real volume at the temperature t of the mercury remaining in the tube is

$$\frac{P-p}{d}(1 + kt).$$

Hence

$$(1 + \kappa t)(1 + k't) = (1 + kt);$$
$$(1 + \kappa t + k't + \kappa k't^2) = 1 + kt.$$

Since κ and k' are small quantities, their product may be neglected; hence

$$k' = k - \kappa. \tag{50}$$

133. **Determination of Voluminal Expansion of Solids.**—The weight thermometer may be used to determine the coefficient of voluminal expansion of solids. For this purpose, the solid, of which the volume at zero is known, must be introduced into the bulb by the glass-blower. If the bulb containing the solid be filled with mercury at zero, and afterward heated to the temperature t, it is evident that the amount of mercury that will overflow will depend upon the coefficient of

expansion of the solid, and upon the coefficient of apparent expansion of mercury. If the latter has been determined for the kind of glass used, the former can be deduced. By this means the coefficients of voluminal expansion of some solids have been determined; and the results are found to verify the conclusion, deduced from theory (§ 128), that the voluminal coefficient is three times the linear.

134. Absolute Expansion of Liquids other than Mercury.—The weight thermometer may also serve to determine the coefficients of expansion of liquids other than mercury; for, if k' has been found as described above, the instrument may be filled with the liquid the coefficient of which is desired, and the apparent expansion of this liquid found exactly as was that of mercury. The absolute coefficient for the liquid is then the sum of the coefficient of apparent expansion and the coefficient for the glass.

135. Expansion of Water.—The use of water as a standard with which to compare the densities of other substances makes it necessary to know, not merely its mean coefficient of expansion, but its actual expansion, degree by degree. This is the more important since water expands very irregularly. The best determinations of the volumes of water at different temperatures are those of Matthiessen. The method which he employed was to weigh in water a mass of glass of which the coefficient of expansion had been previously determined.

Water contracts, instead of expanding, from 0° to 4°; from that temperature to its boiling-point it expands.

136. Correction Introduced in the Determination of Specific Gravity.—Water at its maximum density, at 4°, is the standard to which are referred the specific gravities of solids and liquids. Since it is seldom practicable to make the determinations at that temperature, corrections must be made as follows:

Let Δ_{tt} represent the density of a substance at the temperature t compared to water at the same temperature. Let Δ_t represent the density of the substance at $t°$ compared to water at $4°$. Then $\Delta_t = \Delta_{tt} \times d_t$, where d_t represents the density of water at $t°$; for, if W represent the mass of the substance, W_t the mass of an equal volume of water at $t°$, and W_4 the mass of the same volume of water at $4°$, we have $\Delta_{tt} = \frac{W}{W_t}$, $d_t = \frac{W_t}{W_4}$, and $\Delta_t = \frac{W}{W_4}$. Whence $\Delta_t = \frac{W}{W_t} \times \frac{W_t}{W_4} = \Delta_{tt} \times d_t$. If Δ_0 represent the density of a substance at $0°$, $\Delta_0 = \Delta_t(1 + kt)$, where k represents the coefficient of voluminal expansion of the substance.

137. Effect of Variation of Temperature upon Specific Heat.—It has already been seen (§ 109) that the specific heat of bodies changes with temperature. With most substances the specific heat increases as the temperature rises.

For example, the true specific heat of the diamond

At 0° is	0.0947
At 50° is	0.1435
At 100° is	0.1905
At 200° is	0.2719

138. Effect of Change of Physical State upon Specific Heat.—The specific heat of a substance is not the same when in the solid as when in liquid state. In the solid state of the substance it is generally less than in the liquid. For example:

	Mean Specific Heat	
	Solid.	Liquid.
Water,	0.504	1.000
Mercury,	0.0314	0.333
Tin,	0.056	0.0637
Lead,	0.0314	0.0402

139. Atomic Heat.—It has been found that the product of the specific heat by the atomic weight of any simple body is a constant quantity. This law is known from its discoverers as the *law of Dulong and Petit.*

This law may be otherwise stated, thus: that to raise the temperature of an atom of any simple substance one degree, an amount of heat is required which is the same for all substances.

The experiments of Regnault show that this law may be extended to compound bodies; that is, for all compounds of similar chemical composition the product of the total chemical equivalent by the specific heat is the same.

The following table will illustrate the law of Dulong and Petit. The atomic weights are those given by Clarke.

Elements.	Specific Heat of Equal Weights.	Atomic Weight.	Product of Specific Heat into Atomic Weight.
Iron, ...	0.114	55.9	6.372
Copper, ...	0.095	63.17	6.001
Mercury, . . .	0.0314 (solid)	199.71	6.128
Silver, ...	0.057	107.67	6.137
Gold, ...	0.0329	196.15	6.453
Tin, ...	0.056	117.7	6.591
Lead, ...	0.0314	206.47	6.483
Zinc, ...	0.0955	64.9	6.198

140. Fusion and Solidification.—When ice at a temperature below zero is heated, its temperature rises to zero, and then the ice begins to melt; and, however high the temperature of the medium that surrounds it may be, its temperature remains constant at zero so long as it remains in the solid state. This temperature is the *melting-point of ice*, and because of its fixity it is used as one of the standard temperatures in graduating thermometric scales. Other bodies melt at very different but at fixed and definite temperatures. Many substances cannot be melted, as they decompose by heat.

Alloys often melt at a lower temperature than either of their constituents. An alloy of one part lead, one part tin, four parts bismuth, melts at 94°; while the lowest melting-point of its constituents is that of tln, 228°. An alloy of lead, tin, bismuth, and cadmium melts at 62°.

If a liquid be placed in a medium the temperature of which is below its melting-point, it will, in general, begin to solidify when its temperature reaches its melting-point, and it will remain at that temperature until it is all solidified. Under certain conditions, however, the temperature of a liquid may be lowered several degrees below its melting-point without solidification, as will be seen below.

141. Change of Volume with Change of State.—Substances are generally more dense in the solid than in the liquid state, but there are some notable exceptions. Water, on solidifying, expands; so that the density of ice at zero is only 0.9167, while that of water at 4° is 1. This expansion exerts considerable force, as is evidenced by the bursting of vessels and pipes containing water.

142. Change of Melting- and Freezing-Points.—If water be enclosed in a vessel sufficiently strong to prevent its expansion, it cannot freeze except at a lower temperature. The freezing-point of water is, therefore, lowered by pressure. On the other hand, substances which contract on solidifying have their solidification hastened by pressure.

The lowering of the melting-point of ice by pressure explains some remarkable phenomena. If pieces of ice be pressed together, even in warm water, they will be firmly united. Fragments of ice may be moulded, under heavy pressure, into a solid, transparent mass. This soldering together of masses of ice is called *regelation.* If a loop of wire be placed over a block of ice and weighted, it will cut its way slowly through the ice, and regelation will occur behind it. After the wire has passed through, the block will be found one solid mass, as before. The explanation of these phenomena is, that the ice is partially melted by the pressure. The liquid thus formed is colder than the ice; it finds its way to points of less pressure, and there, because of its low temperature, it congeals, firmly uniting the two masses.

Water, when freed from air and kept perfectly quiet, will not form ice at the ordinary freezing-point. Its temperature may be lowered to -10° or -12° without solidification. In this condition a slight jar, or the introduction of a small fragment of ice, will cause a sudden congelation of part of the liquid, accompanied by a rise in temperature in the whole mass to zero.

A similar phenomenon is observed in the case of several solutions, notably sodium sulphate and sodium acetate. If a saturated hot solution of one of these salts be made, and allowed to cool in a closed bottle in perfect quiet, it will not crystallize. Upon opening the bottle and admitting air, crystallization commences, and spreads rapidly through the mass, accompanied by a considerable rise of temperature. If the amount of salt dissolved in the water be not too great, the solution will remain liquid when cooled in the open air, and it may even suffer considerable disturbance by foreign bodies without crystallization; but crystallization begins immediately upon contact with the smallest crystal of the same salt.

143. Heat Equivalent of Fusion.—Some facts that have appeared in the above account of the phenomena of fusion and solidification require further study. It has been seen that, however rapidly the temperature of a solid may be rising, the moment fusion begins the rise of temperature ceases. Whatever the heat to which a solid may be exposed, it cannot be made hotter than its melting-point. When ice is melted by pressure, its temperature is lowered. When a liquid is cooled, its fall of temperature ceases when solidification begins; and if, as may occur under favorable conditions, a liquid is cooled below its melting-point, its temperature rises at once to the melting-point, when solidification begins. Heat, therefore, disappears when a body melts, and is generated when a liquid becomes solid.

It was stated (§ 101) that ice can be melted by friction; that is, by the expenditure of mechanical energy. Fusion is, therefore, work which requires the expenditure of some form of energy to accomplish it. The heat required to melt unit mass of a substance is the *heat equivalent of fusion* of that substance. When a substance solidifies, it develops the same amount of heat as was required to melt it.

144. Nature of the Energy stored in the Liquid.—From the facts given above, as well as from the principle of the conservation of energy, it appears that the energy expended in melting a body is stored in the liquid. It is easy to see what must be the nature of this energy. When a body solidifies, its molecules assume certain positions in obedience to their mutual attractions. When it is melted, the molecules are forced into new positions in opposition to the attractive forces. They are, therefore, in positions of advantage with respect to these forces, and possess potential energy.

145. Determination of the Heat Equivalent of Fusion.—The heat equivalent of fusion may be determined by the method of mixtures (§ 111), as follows: a mass of ice, for example, represented by P, at a temperature t below its melting-point, to insure dryness, is plunged into a mass P' of warm water at the temperature T. Represent by θ the resulting temperature, when the ice is all melted. If p represent the water equivalent of the calorimeter, $(P' + p)(T - \theta)$ is the heat given up by the calorimeter and its contents. Let c represent the specific heat of ice, and x the heat equivalent of fusion. The ice absorbs, to raise its temperature to zero, Ptc calories; to melt it, Px calories; to warm the water after melting, $P\theta$ calories. We then have the equation

$$Ptc + P\theta + Px = (P' + p)(T - \theta),$$

from which x may be found.

Other calorimetric methods may be employed. The best experiments give, for the heat equivalent of fusion of ice, very nearly eighty calories.

GASES AND VAPORS.

146. **The Gaseous State.**—A *gas* may be defined as a highly compressible fluid. A given mass of gas has no definite volume. Its volume varies with every change in the external pressure to which it is exposed. A *vapor* is the gaseous state of a substance which at ordinary temperatures exists as a solid or a liquid.

147. **Vaporization** is the process of formation of vapor. There are two phases of the process, *evaporation*, in which vapor is formed at the free surface of the liquid, and *ebullition*, in which the vapor is formed in bubbles in the mass of the liquid, or at the heated surface with which it is in contact.

148. **Nature of the Process of Evaporation.**—It has been seen (§ 101) that there are many reasons for believing that the molecules of solids and liquids are in a state of continual motion. It is not supposed that any one molecule maintains continuously the same condition of motion; but in the interaction of the molecules the motion of any one may be more or less violent, as it receives motion from its neighbors, or gives up motion to them. It can easily be supposed that, at the exposed surface of the substance, the motion of a molecule may at times be so violent as to project it beyond the reach of the molecular attractions. If this occur in the air, or in a space filled with any gas, the molecule may be turned back, and made to rejoin the molecules in the liquid mass; but many will find their way to such a distance that they will not return. They then constitute a vapor of the substance. As the number of free molecules in the space above the liquid increases, it is plain that there may come a time when as many will rejoin the liquid as escape from it. The space is then *saturated* with the vapor.

The more violent the motion in the liquid, that is the higher its temperature, the more rapidly the molecules will escape, and the greater must be the number in the space above the liquid before the returning will equal in number the outgoing molecules. In other words, the higher the temperature, the more dense the vapor that saturates a given space. If the space above a liquid be a vacuum, the escaping molecules will at first meet with no obstruction, and, as a consequence, the space will be very quickly saturated with the vapor.

Experiment verifies all these deductions. Evaporation goes on continually from the free surfaces of many liquids, and even of solids. It increases in rapidity as the temperature increases, and ceases when the vapor has reached a certain density, always the same for the same temperature, but greater for a higher temperature. It goes on very rapidly in a vacuum; but it is found that the final density of the vapor is no greater, or but little greater, than when some other gas is present. In other words, while a foreign gas impedes the motion of the outgoing molecules, and causes evaporation to go on slowly, it has very little influence upon the number of molecules that must be present in order that those which return may equal in number those which escape.

149. **Pressure of Vapors.**—As a liquid evaporates in a closed space, the vapor formed exerts a pressure upon the enclosure and upon the surface of the liquid, which increases so long as the quantity of vapor increases, and reaches a maximum when the space is saturated. This *maximum pressure* of a vapor increases with the temperature. When evaporation takes place in a space filled by another gas which has no action upon the vapor, the pressure of the vapor is added to that of the gas, and the pressure of the mixture is, therefore, the sum of the pressures of its constituents. The law was announced by Dalton that the quantity of vapor which saturates a given space, and consequently the maximum pressure of that vapor, is the same whether the space be empty or contain a gas.

Regnault has shown that, for water, ether, and some other substances, the maximum pressure of their vapors is slightly less when air is present.

150. Ebullition.—As the temperature of a liquid rises, the pressure which its vapor may exert increases, until a point is reached where the vapor is capable of forming, in the mass of the liquid, bubbles which can withstand the superincumbent pressure of the liquid and the atmosphere above it. These bubbles of vapor, escaping from the liquid, give rise to the phenomenon called *ebullition*, or *boiling*. Boiling may, therefore, be defined as the agitation of a liquid by its own vapor.

Generally speaking, for a given liquid, ebullition always occurs at the same temperature for the same pressure; and, when once commenced, the temperature of the liquid no longer rises, no matter how intense the source of heat. This fixed temperature is called the *boiling-point* of the liquid. It differs for different liquids, and for the same liquid under different pressures. That the boiling-point must depend upon the pressure is evident from the explanation of the phenomenon of ebullition above given.

Substances in solution, if less volatile than the liquid, retard ebullition. While pure water boils at 100°, water saturated with common salt boils at 109°. The material of the containing vessel also influences the boiling point. In a glass vessel the temperature of boiling water is higher than in one of metal. If water be deprived of air by long boiling, and then cooled, its temperature may afterwards be raised considerably above the boiling-point before ebullition commences. Under these conditions, the first bubbles of vapor will form with explosive violence. The air dissolved in water separates from it at a high temperature in minute bubbles. Into these the water evaporates, and, whenever the elastic force of the vapor is sufficient to overcome the superincumbent pressure, it enlarges them, and causes the commotion that marks the phenomenon of ebullition. If no such openings in the mass of the fluid exist, the cohesion of the fluid, or its adhesion to the vessel, as well as the pressure, must be overcome by the vapor. This explains the higher temperature at which ebullition commences when the liquid has been deprived of air.

151. Spheroidal State.—If a liquid be introduced into a highly heated capsule, or poured upon a very hot plate, it does not wet the heated surface, but forms a flattened spheroid, which presents no appearance of boiling, and evaporates only very slowly. Boutigny has carefully studied these phenomena, and made known the following facts. The temperature of the spheroid is below the boiling-point of the liquid. The spheroid does not touch the heated plate, but is separated from it by a non-conducting layer of vapor. This accounts for the slowness of the evaporation. To maintain the liquid in this condition the temperature of the capsule must be much above the boiling-point of the liquid; for water it must be at least 200° C. If the capsule be allowed to cool, the temperature will soon fall below the limit necessary to maintain the spheroidal state, the liquid will moisten the capsule, and there will be a rapid ebullition, with disengagement of vapor. If a liquid of very low boiling-point, as liquid nitrous oxide, which boils at — 88°, be poured into a red-hot capsule, it will assume the spheroidal state; and, since its temperature cannot rise above its boiling-point, water, or even mercury, plunged into it, will be frozen.

152. Production of Vapor in a Limited Space.—When a liquid is heated in a limited space the vapor generated accumulates, increasing the pressure, and the temperature rises above the ordinary boiling-point. Cagniard-Latour experimented upon liquids in spaces but little larger than their own volumes. He found that, at a certain temperature, the liquid suddenly disappeared; that is, it was converted into vapor in a space but little larger than its own volume. It is supposed that above the temperature at which this occurs, which is called the *critical temperature,* the substance cannot exist in the liquid state.

153. Liquefaction.—Only a certain amount of vapor can exist at a given temperature in a given space. If the temperature of a space saturated with vapor be lowered, some of the vapor must condense into the liquid state. It is not necessary that the temperature of the whole space be lowered; for, when the vapor in the cooled portion is condensed, its pressure is diminished, the vapor from the warmer portion flows in, to be in its turn condensed, and this continues until the whole is brought to the density and pressure due to the cooled portion. Any diminution of the space occupied by a saturated vapor at constant temperature, will cause some of the vapor to become liquid, for, if it do not condense, its pressure must increase; but a saturated vapor is already at its maximum pressure.

If the vapor in a given space be not at its maximum pressure, its pressure will increase when its volume is diminished; until the maximum pressure is reached; when, if the temperature remain constant, further reduction of volume causes condensation into the liquid state, without further increase of pressure or density. This statement is true of several of the gases at ordinary temperatures. Chlorine, sulphur dioxide, ammonia, nitrous oxide, carbon dioxide, and several other gases, become liquid under sufficient pressure. Andrews found that, at a temperature of 30.92°, pressure ceases to liquefy carbon dioxide. This is the critical temperature for that substance. The critical temperatures of oxygen, hydrogen, and the other so-called permanent gases, are so low that it is only by methods capable of yielding an extremely low temperature that they can be liquefied. By the use of such methods any of the gases may be made to assume the liquid state. In the case of hydrogen, however, the low temperature necessary for its liquefaction has only been reached by allowing the gas to expand suddenly from a condition of great condensation, in which it had already been cooled to a very low point.

154. Pressure and Density of Non-saturated Gases and Vapors.—If a gas or vapor in the non-saturated condition be maintained at constant temperature, it follows very nearly Boyle's law (§§ 76 and 98). If its temperature be below its critical temperature, the product of volume by pressure diminishes, and near the point of saturation the departure from the law may be considerable. At this point there is a sudden diminution of volume, and the gas assumes the liquid state. The less the pressure and density of the gas, the more nearly it obeys Boyle's law.

It has been stated already (§ 99) that gases expand as the temperature rises. The law of this expansion, called, after its discoverer, *Gay-Lussac's law*, is that, for each increment of temperature of one degree, every gas expands by the same constant fraction of its volume at zero. This is equivalent to saying that a gas has a constant coefficient of expansion, which is the same for all gases.

Let V_0, V_t, represent the volumes at zero and t respectively, and α the coefficient of expansion. Then, the pressure remaining constant, we have

$$V_t = V_0(1 + \alpha t). \tag{51}$$

If d_0, d_t, represents the densities at the same two temperatures, we have, since densities are inversely as volumes,

$$d_t = \frac{d_0}{1 + \alpha t}. \tag{52}$$

Later investigations, especially those of Regnault, show that this simple law, like the law of Boyle, is not rigorously true, though it is very nearly so for all gases and vapors which are not too near their points of saturation. The common coefficient of expansion is $\alpha = 0.003666 = \frac{1}{273}$ very nearly.

From the law of Boyle we have, for a given mass of gas, if the temperature remain constant,

$$V_p p = V_{p'} p' = \text{volume at pressure unity},$$

where V_p, $V_{p'}$, represent the volumes at pressure p and p' respectively.

From the law of Gay-Lussac we have, if the pressure remain constant,

$$V_0 = \frac{V_t}{1 + \alpha t} = \frac{V_{t'}}{1 + \alpha t'}. \tag{53}$$

If the temperature and pressure both vary, we have

$$\frac{V_{pt} p}{1 + \alpha t} = \frac{V_{p't'} p'}{1 + \alpha t'}; \tag{54}$$

that is, if the volume of a given mass of gas be multiplied by the corresponding pressure and divided by the factor of expansion, the quotient is constant.

155. **Pressure and Density of Saturated Gases and Vapors.**—It has been seen that, for each gas or vapor at a temperature below the critical temperature, there is a maximum pressure which it can exert at that temperature. To each temperature there corresponds a maximum pressure, which is higher as the temperature is higher. A gas or vapor in contact with its liquid in a closed space will exert its maximum pressure.

The relation between the temperature and the corresponding maximum pressure of a vapor is a very important one, and has been the subject of many investigations. The vapor of water has been especially studied, the most extensive and accurate experiments being those of Regnault.

Two distinct methods were employed, one for temperatures below 50°, and the other for higher temperatures. The first consisted in observing the difference in height of two barometers placed side by side, the vacuum chamber of one containing a little water. The temperature was carried from zero to about 50°. Both barometers were surrounded by the same medium, and in every way under the same conditions, except that water and its vapor were present in one and not in the other. The difference between the heights of the two gave the pressure of the vapor at the temperature of the experiment.

The second method was founded on the principle that the vapor of a boiling liquid exerts a pressure equal to that of the atmosphere above it. The experiment consisted in boiling water in a closed space in which the air could be rarefied or condensed to a known pressure, and noting the temperature of the boiling liquid and that of the vapor above it. To prevent the accumulation of the vapor and the consequent change of pressure, a condenser communicated with the boiler, consisting of a tube surrounded by a larger tube, forming an annular space, through which a stream of cold water was kept flowing. By this means the vapor was condensed as fast as formed, and the water from its condensation flowed back into the boiler. By rarefying or compressing the air in the closed space, an artificial atmosphere of any desired pressure could be obtained, and maintained constant as long as was necessary for making the observations.

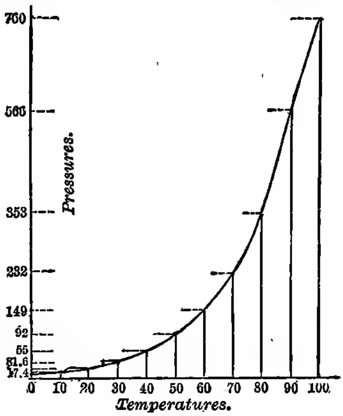

FIG. 51.

The temperature was determined by means of four thermometers placed in the boiler, two of them in the liquid and two in the vapor. The bulbs of the thermometers were placed in metal tubes, to protect them from the pressure, which otherwise would compress the bulb, and cause the thermometer to register too high a temperature.

The results of Regnault's observations may be represented graphically, as in Fig. 51, where pressures are measured in the vertical, and temperatures in the horizontal, direction. It is seen that the pressure varies very rapidly with the temperature.

156. Kinetic Theory of Gases.—According to the *kinetic theory of gases*, a perfect gas consists of an assemblage of free, perfectly elastic molecules in constant motion. Each molecule moves in a straight line with a constant velocity, until it encounters some other molecule, or the side of the vessel. The impacts of the molecules upon the sides of the vessel are so numerous that their effect is that of a continuous constant force or pressure.

The entire independence of the molecules is assumed from the fact that, when gases or vapors are mixed, the pressure of one is added to that of the others; that is, the pressure of the mixture is the sum of the pressures of the separate gases. It follows from this, that no energy is required to separate the molecules; in other words, no internal work need be done to expand a gas. This was demonstrated experimentally by Joule, who showed that when a gas expands without performing external work, it is not cooled.

The action between two molecules, or between a molecule and a solid wall, must be of such a nature that no energy is lost; that is, the sum of the kinetic energies of all the molecules must remain constant. Whatever be the nature of this action, it is evident that when a molecule strikes a solid stationary wall, it must be reflected back with a velocity equal to that before impact. If the velocity be resolved into two components, one parallel to the wall and the other normal to it, the parallel component remains unchanged, while the normal component is changed from $+v$, its value before impact, to $-v$, its value after impact. The change of velocity is therefore $2v$; and if θ represent the duration of impact, the mean acceleration is $\frac{2v}{\theta}$, and the mean force of impact $p = m\frac{2v}{\theta}$, where m represents the mass of the molecule.

Since the effect of the impacts is a continuous pressure, the total pressure P exerted upon unit area is equal to this mean force of impact of one molecule multiplied by the number of molecules meeting unit area in the time θ. To find this latter factor, we suppose the molecules confined between two parallel walls at a distance s from each other. Any molecule may be supposed to suffer reflection from one wall, pass across to the other, be reflected back to the first, and so on. Whatever may be the effect of the mutual collisions of the molecules, the number of impacts upon the surface considered will be the same as though each one preserved its rectilinear motion unchanged, except when reflected from the solid walls. The time required for a molecule moving with a velocity v to pass across the space between the two walls and back is $\frac{2s}{v}$; and the number of impacts upon the first surface in unit time is $\frac{v}{2s}$.

Represent by n' the number of molecules in a rectangular prism, with bases of unit area in the walls. These molecules must be considered as moving in all directions and with various velocities. But the velocity of any molecule may be resolved in the direction of three rectangular axes, one normal to the surface and the other two parallel to it; and, since the number of molecules in any finite volume of gas is practically infinite, the effect upon the wall due to their real motions will be the same as would result from a motion of one third the total number of molecules in each of the three directions with the mean velocity. Hence the number of molecules moving, in a manner similar to that of the single molecule already considered, normal to the walls is $\frac{1}{3}n'$. The number of impacts upon unit area of the first surface in unit time is $\frac{1}{3}\frac{n'v}{2s}$; and in time θ is $\frac{1}{3}\frac{n'v\theta}{2s}$. Hence the total pressure P on unit area is

$$P = m\frac{2v}{\theta} \times \frac{1}{3}\frac{n'v\theta}{2s} = \frac{1}{3}mv^2\frac{n'}{s}.$$

But $\frac{n'}{s}$ is the number of molecules in unit volume. Representing this by n, we have

$$P = \tfrac{1}{3}nmv^2. \tag{55}$$

That is, the pressure upon unit area is equal to one third the number of molecules in unit volume at that pressure multiplied by twice the kinetic energy of each molecule.

Suppose, now, the volume of the gas be changed from unity to V, without change of temperature. The number of molecules in unit volume is now $\frac{n}{V}$, and the pressure $P_1 = \frac{1}{3}\frac{n}{V}mv^2$, whence $P_1V = \frac{1}{3}nmv^2$. This is a constant quantity, since n and m are constant for the same mass of gas, and v is constant if there be no change of temperature. But PV equal to a constant is Boyle's law.

From the law of Gay-Lussac we have, if P represent the pressure at $t°$, and P_0 the pressure at zero,

$$P = P_0(1 + \alpha t).$$

We have $\alpha = \frac{1}{273}$ very nearly; hence

$$P = P_0\left(1 + \frac{t}{273}\right). \tag{56}$$

If $t = -273°$,

$$P = P_0\left(1 - \frac{273}{273}\right) = 0;$$

that is, at $273°$ below zero the pressure vanishes. Since $P = \frac{1}{3}nmv^2$, it follows that, at this temperature, $v = 0$, or the molecules are at rest. This temperature is therefore called the *absolute zero*.

In studying the expansion of gases, it is very convenient to use a scale of temperatures the zero-point of which is at the absolute zero. Temperatures reckoned upon this scale are called absolute temperatures. Let T represent a temperature upon the absolute scale: then $T = t + 273$, and Eq. (56) becomes $P = P_0\frac{T}{273}$. Substituting the value of P from (55), we have

$$\tfrac{1}{3}nmv^2 = P_0\frac{T}{273},$$

whence

$$T = \tfrac{1}{3}n\frac{273}{P_0}mv^2. \tag{57}$$

That is, the absolute temperature of a gas is proportional to the kinetic energy of the molecules.

It has been already stated (§ 100), that, when a gas is compressed, a certain amount of heat is generated. Suppose a cylinder with a tightly-fitting piston. So long as the piston is at rest, each molecule that strikes it is reflected with a velocity equal to that before impact: but if the piston be forced into the cylinder, each molecule, as it is reflected, has its velocity increased; and, as was shown above, this is equivalent to a rise in temperature. It can be shown that the increase of kinetic energy in this case is precisely equal to the work done in forcing the piston into the cylinder against the pressure of the gas. On the other hand, if the piston be pushed backward by the force of the impact of the molecules, there will be a loss of velocity by reflection from the moving surface, kinetic energy equal in amount to the work done upon the piston disappears, and the temperature falls.

The phenomena exhibited by the *radiometer* afford a strong experimental confirmation of the kinetic theory of gases. These phenomena were discovered by Crookes. In the form first given to it by him, the instrument consists of a delicate torsion balance suspended in a vessel from which the air is very completely exhausted. On one end of the arm of the torsion balance is fixed a light vane, one face of which is blackened. When a beam of light falls on the vane, it moves as if a pressure were applied to its blackened surface. The explanation of this movement is, that the molecules of air remaining in the vessel are more heated when they come in contact with the blackened face of the vane than when they come in contact with the other face, and are hence thrown off with a greater velocity, and react more strongly upon the blackened face of the vane. At ordinary pressures the free paths of the molecules are very small, their collisions very frequent, and any inequality in the pressures is so speedily reduced, that no effect upon the vane is apparent. At the high exhaustions at which the movement of the vane becomes evident, the collisions are less frequent, and hence an immediate equalization of pressure does not occur. The vane therefore moves in consequence of the greater reaction upon its blackened surface.

157. Mean Velocity of Molecules.—Equation (55) enables us to determine the mean velocity of the molecules of a gas of which the density and pressure are known, since nm is the mass of the gas in unit volume.

Solving the equation with reference to v, and substituting the known values of the constants for hydrogen, namely, $P = 1013373$ dynes per square centimetre, and nm, or density, $= 0.00008954$ grams per cubic centimetre, we have 184260 centimetres per second, or a little more than one mile per second, as the mean velocity of a molecule of hydrogen.

158. Elasticity of Gases.—It has been shown (§ 77) that the elasticity of a gas, obeying Boyle's law, is numerically equal to the pressure. This is the *elasticity for constant temperature.* But, as was seen (§ 156), when a gas is compressed it is heated; and heating a gas increases its pressure. Under ordinary conditions, therefore, the ratio of a small increase of pressure to the corresponding decrease of unit volume is greater than when the temperature is constant. It is important to consider the case when all the heat generated by the compression is retained by the gas. The elasticity is then a maximum, and is called the *elasticity when no heat is allowed to enter or escape.*

Let mn (Fig. 52) be a curve representing the relation between volume and pressure for constant temperature, of which the abscissas represent volumes and the ordinates pressures. Such a curve is called an *isothermal* line. It is plain that to each temperature must correspond its own isothermal line. If, now, we suppose the gas to be compressed, and no heat to escape, it is plain that if the volume diminish from OC to OG, the pressure will become greater than GD; suppose it to be GM. If a number

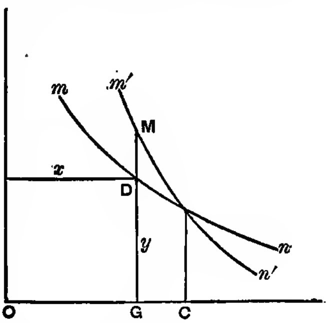

FIG. 52.

of such points as M be found, and a line be drawn through them, it will represent the relation between volume and pressure when no heat enters or escapes. It is called an *adiabatic* line. It evidently makes a greater angle with the horizontal than the isothermal.

159. Specific Heats of Gases.—In § 156 it is seen that the temperature of a gas is proportional to the kinetic energy of its molecules. To warm a gas without change of volume is, therefore, only to add to this kinetic energy. If, however, the gas be allowed to expand when heated, the molecules lose energy by impact upon the receding surface; and this, together with the kinetic energy due to the rise in temperature, must be supplied from the source of heat. It has been seen that the loss of energy resulting from impact upon a receding surface is equal to the work done by the gas in expanding.

The amount of heat necessary to raise the temperature of unit mass of a gas one degree, while the volume remains unchanged, is called *the specific heat of the gas at constant volume.* The amount of heat necessary to raise the temperature of unit mass of a gas one degree when expansion takes place without change of pressure, is called *the specific heat of the gas at constant pressure.*

From what has been said above, it is evident that the difference between these two quantities of heat is the equivalent of the work done by the expanding gas.

The determination of the relation of these two quantities is a very important problem.

The specific heat of a gas at constant pressure may be found by passing a current of warmed gas through a tube coiled in a calorimeter. This is the method of mixtures (§ 111). There are great difficulties in the way of an accurate determination, because of the small density of the gas, and the time required to pass enough of it through the calorimeter to obtain a reasonable rise of temperature. The various sources of error produce effects which are sometimes as great as, or even greater than, the quantity to be measured. It is beyond the scope of this work to describe in detail the means by which the effects of the disturbing causes have been determined or eliminated.

The specific heat of a gas at constant volume is generally determined from the ratio between it and the specific heat at constant pressure. The first determination of this ratio was accomplished by Clement and Desormes.

The theory of the experiment may be understood from the following considerations:

Let a unit mass of gas at any temperature t and volume V_t be confined in a cylinder by a closely fitting piston of area A. Suppose its temperature to be raised one degree, by communication of heat from some external source, while its volume remains unchanged. It absorbs heat, which we will suppose measured in mechanical units, and will represent by C_v the specific heat at constant volume. Now let the gas expand, at the constant temperature $t + 1$, until it returns to its original pressure. During this expansion the piston will be forced out through a distance d, and an additional quantity of heat will be absorbed from the source. Represent by P the mean pressure on unit area of the piston exerted by the gas during this operation. Then the work done during expansion, which is the equivalent of the heat absorbed, is PAd. Ad represents the increase in volume of the gas during this process. The same increase in volume would have occurred had the gas been allowed to expand at constant pressure, while its temperature was rising. But, for a rise in temperature of one degree, the increase in volume of any mass of gas is αV_0, where V_0 represents the volume at zero. Hence we have $Ad = \alpha V_0$, and the work done during the expansion is $PAd = P\alpha V_0$. The heat absorbed, therefore, in raising the temperature of the gas one degree at constant pressure is $C_p = C_v + P\alpha V_0$. C_p represents the specific heat of the gas at constant pressure, measured in mechanical units. The ratio of the two specific heats is

$$\frac{C_p}{C_v} = 1 + \frac{1}{C_v} P\alpha V_0. \tag{58}$$

If, in the case considered above, the gas had expanded without receiving any heat, the work $P\alpha V_0$ would have been done at the expense of its own internal energy, and the temperature would have fallen. The performance of this work is equivalent to abstracting the quantity of heat, $P\alpha V_0$, which would lower the temperature $\frac{1}{C_v} \cdot P\alpha V_0$ degrees, since the abstraction of a quantity C_v of heat would lower the temperature one degree. Represent this change of temperature by θ. Remembering that the supposed change of volume was αV_0, which equals $\frac{\alpha V_t}{1 + \alpha t}$, and that the original volume was V_t, it is seen that the change of $\frac{\alpha}{1 + \alpha t}$ in unit volume would cause a fall in temperature of θ degrees. Substituting θ for $\frac{1}{C_v} P\alpha V_0$ in Eq. (58), we have $\frac{C_p}{C_v} = 1 + \theta$. It is the object of the experiment to find θ. The method of Clement and Desormes is as follows:

A large flask is furnished with a stopcock having a large opening, and a very sensitive manometer which shows the difference between the pressure in the flask and the pressure of the air. The air in the flask is first rarefied, and left to assume the temperature of the surrounding atmosphere. Suppose its pressure now to be $H - h$, H representing the height of the barometer, and h the difference between the pressure in the flask and the pressure of the atmosphere, as shown by the manometer. The large stopcock is then suddenly opened for a very short time only; the air rushes in, re-establishes the atmospheric pressure, compresses the air originally in the flask, and raises its temperature. The volume of the air becomes $1 - \phi$, where its original volume is taken as unity and ϕ represents its reduction; and, if there were no change of temperature, the pressure would be $\frac{H-h}{1-\phi}$. If the temperature increase θ' degrees, and become $t + \theta'$, the pressure will be

$$\frac{H-h}{1-\phi} \times \frac{1+\alpha(t+\theta')}{1+\alpha t} = H, \tag{59}$$

the atmospheric pressure.

The flask is now left until the air within it returns to the temperature of the atmosphere t, when the manometer shows a fall of pressure h', and we have

$$\frac{H-h}{1-\phi} = H - h'. \tag{60}$$

From these two equations we have

$$\phi = \frac{h-h'}{H-h'}; \quad \theta' = \frac{(1+\alpha t)h'}{\alpha(H-h')}.$$

Suppose, now, the change of volume had been $\frac{\alpha}{1+\alpha t}$, then the change of temperature would have been θ; and, since change of volume is proportional to change of temperature, we have

$$\phi : \frac{\alpha}{1+\alpha t} = \theta' : \theta;$$

hence

$$\theta = \frac{\theta' \frac{\alpha}{1+\alpha t}}{\phi};$$

or, substituting the values of ϕ and θ', we have

$$\theta = \frac{h'}{H - h'} \times \frac{H - h'}{h - h'} = \frac{h'}{h - h'}.$$

Now we have shown that

$$\frac{C_p}{C_v} = 1 + \theta;$$

hence

$$\frac{C_p}{C_v} = 1 + \frac{h'}{h - h'} = \frac{h}{h - h'}. \tag{61}$$

160. The Two Specific Heats of a Gas have the Same Ratio as the Two Elasticities.—Suppose a gas, of which the mass is unity and volume V, to rise in temperature at constant pressure from the temperature t to the temperature $(t + \Delta t)$, Δt representing a very small increment of temperature. The heat consumed will be $C_p \Delta t$, and the increase of volume $\alpha V_0 \Delta t$. Now, if the volume had remained constant, the amount of heat required to cause the rise of temperature Δt would have been $C\ \Delta t$. Hence if the gas be not allowed to expand, the amount of heat, $C_p \Delta t$, will cause a rise of temperature $\frac{C_p}{C_v} \Delta t$; and the same rise of temperature will occur if the gas, after first being allowed to expand, be compressed to its initial volume. Such a compression would be attended by an increase of pressure, which we will call Δp. The ratio between this and the corresponding change of volume is

$$\frac{\Delta p}{\alpha V_0 \Delta t} = E_h, \tag{62}$$

where E_h is the elasticity under the condition that no heat enters or escapes.

If, now, the heat produced by compression be allowed to escape, there will remain the quantity $C_v \Delta t$, and the increment of pressure will be reduced to $\delta p = \Delta p \frac{C_v}{C_p}$. This is the increase of pressure that will occur if the gas be compressed by the amount $\alpha V_0 \Delta t$ without change of temperature; hence

$$\frac{\delta p}{\alpha V_0 \Delta t} = E_t, \tag{63}$$

where E_t is the elasticity for constant temperature. Dividing (62) by (63), we have

$$\frac{E_h}{E_t} = \frac{\dfrac{\Delta p}{\alpha V_0 \Delta t}}{\dfrac{\delta p}{\alpha V_0 \Delta t}} = \frac{\Delta p}{\delta p} = \frac{\Delta p}{\Delta p \dfrac{C_v}{C_p}} = \frac{C_p}{C_v};$$

that is, the two elasticities have the same ratio as the two specific heats of a gas.

It may be shown that the velocity of sound in any medium is equal to the square root of the quotient of the elasticity divided by the density of the medium; that is,

$$velocity = \sqrt{\frac{E}{D}}. \tag{64}$$

In the progress of a sound-wave, the air is alternately compressed and rarefied, the compressions and rarefactions occurring in such rapid succession that there is no time for any transfer of heat. If Eq. (64) be applied to air, the E becomes E_h, or the elasticity under the condition that no heat enters or escapes. Since we know the density of the air and the velocity of sound, E_h can be computed. In § 77 it is shown that E_t is numerically equal to the pressure; hence we have the values of the two elasticities of air, and, as seen above, their ratio is the ratio of the two specific heats of air.

161. Examples of Energy absorbed by Vaporization.—When a liquid boils, its temperature remains constant, however intense the source of heat. This shows that the heat applied to it is expended in producing the change of state. Heat is absorbed during evaporation. By promoting evaporation, intense cold may be produced. In a vacuum, water may be frozen by its own evaporation. If a liquid be heated to a temperature above its ordinary boiling-point under pressure, relief of the pressure is followed by a very rapid evolution of vapor and a rapid cooling of the liquid. Liquid nitrous oxide at a temperature of zero is still far above its boiling-point, and its vapor exerts a pressure of about thirty atmospheres. If the liquid be drawn off into an open vessel, it at first boils with extreme violence, but is soon cooled to its boiling-point for the atmospheric pressure, about $-88°$, and then boils away slowly, while its temperature remains at that low point.

162. Heat Equivalent of Vaporization.—It is plain, from what has preceded (§ 148), that the formation of vapor is work requiring the expenditure of energy for its accomplishment. Each molecule that is shot off into space obtains the motion which projected it beyond the reach of the molecular attraction, at the expense of the energy of the molecules that remain behind. A quantity of heat disappears when a liquid evaporates; and experiment demonstrates, that to evaporate a kilogram of a liquid at a given temperature always requires the same amount of heat. This is the *heat equivalent of vaporization*. When a vapor condenses into the liquid state, the same amount of heat is generated as disappears when the liquid assumes the state of vapor. The heat equivalent of vaporization is determined by passing the vapor at a known temperature into a calorimeter, there condensing it into the liquid state, and noting the rise of temperature in the calorimeter. This, it will be seen, is essentially the method of mixtures. Many experimenters have given attention to this determination; but here, again, the best experiments are those of Regnault. He determined what he called the *total heat of steam* at various pressures. By this was meant the heat required to raise the temperature of a kilogram of water from zero to the temperature of saturated vapor at the pressure chosen, and then convert it wholly into steam. The result of his experiments give, for the heat equivalent of vaporization of water at 100°, 537 calories. That is, he found, that by condensing a kilogram of steam at 100° into water, and then cooling the water to zero, 637 calories were obtained. But almost exactly 100 calories are derived from the water cooling from 100° to zero; hence 537 calories is the heat equivalent of vaporization at 100°.

163. Dissociation.—It has already been noted (§ 99), that, at high temperatures, compounds are separated into their elements. To effect this separation, the powerful forces of chemical affinity must be overcome, and a considerable amount of energy must be consumed.

164. Heat Equivalent of Dissociation and Chemical Union.—From the principle of the conservation of energy, it may be assumed that the energy required for dissociation is the same as that developed by the reunion of the elements. The heat equivalent of chemical union is not easy to determine because the process is usually complicated by changes of physical state. We may cause the union of carbon and oxygen in a calorimeter, and, bringing the products of combustion to the temperature of the elements before the union, measure the heat given to the instrument; but the carbon has changed its state from a solid to a gas, and some of the chemical energy must have been consumed in that process. The heat measured is the *available* heat. The best determinations of the available heat of chemical union have been made by Andrews, Favre and Silbermann, and Berthelot.

HYGROMETRY.

165. Object of Hygrometry.—Hygrometry has for its object the determination of the state of the air with regard to moisture.

The amount of vapor in a given volume of air may be determined directly by passing a known volume of air through tubes containing some substance which will absorb the moisture, and finding the increase in weight of the tubes and their contents. The quantity of vapor contained in a cubic metre of air is called its *absolute humidity*. Methods of determining this quantity indirectly are given below.

166. Pressure of the Vapor.—It has been seen (§ 149), that, when two or more gases occupy the same space, each exerts its own pressure independently of the others. The pressure of the atmosphere is, therefore, the pressure of the dry air, with that of the vapor of water added. If we can determine this latter pressure it is easy to compute the quantity of moisture in the air.

It has also been seen that the pressure exerted by the vapor in the air is at a certain temperature its maximum pressure. Now, if any small portion of the space be cooled till its temperature is below that at which the pressure exerted is the maximum pressure, a portion of the vapor will condense into liquid. If, then, we determine the temperature at which condensation begins, the maximum pressure of the vapor for this temperature, which may be found from tables, is the real pressure of the vapor in the air. The mass of vapor in a cubic metre of air may then be computed as follows: A cubic metre of dry air has a mass of 1293.2 grams at zero and at 700 millimetres pressure. At the pressure p of the vapor, and temperature t of the air at the time of the experiment, the same space would contain

$$1293.2 \times \frac{p}{760} \times \frac{1}{1 + \alpha t}$$

grams of air; and, since the density of vapor of water referred to air is 0.623, a cubic metre would contain

$$1293.2 \times \frac{p}{760} \times \frac{1}{1 + \alpha t} \times 0.623 \qquad (65)$$

grams of vapor.

167. Dew Point.—The temperature at which the vapor of the air begins to condense is called the *dew point*. It is determined by means of instruments called dew-point hygrometers, which are instruments so constructed that a small surface exposed to the air may be cooled until moisture deposits upon it, when its temperature is accurately determined. The Alluard hygrometer consists of a metal box about one and a half centimetres square and four centimetres deep. Two tubes pass through the top of the box—one terminating just inside and the other extending to the bottom. One side of the box is gilded and polished, and is so placed that the gilded surface lies on the same plane with, and in close proximity to, a gilded metal plate. The box is partly filled with ether, and the short tube is connected with an aspirator. Air is thus drawn through the longer tube, and, bubbling up through the ether, causes rapid evaporation, which soon cools the box, and causes a deposit of dew upon the gilded surface. The presence of the gilded plate helps very much in recognizing the beginning of the deposit of dew, by the contrast between it and the dew-covered surface of the box. A thermometer plunged in the ether gives its temperature, and another outside gives the temperature of the air. The temperature of the ether is the dew point. From it the pressure of the vapor in the air is determined as described in the last section, and this pressure substituted for p in Eq. 66 gives the absolute humidity.

168. Relative Humidity.—The amount of moisture that the air may contain depends upon its temperature. The dampness or dryness of the air does not depend upon the absolute amount of moisture it contains, but upon the ratio of this to the amount it might contain if saturated. The *relative humidity* is the ratio of the amount of moisture in the air to that which would be required to saturate it at the existing temperature. Since non-saturated vapors follow Boyle's law very closely, this ratio will be very nearly the ratio of the actual pressure to the possible pressure for the temperature. Both these pressures may be taken from the tables. One corresponds to the dew point, and the other to the temperature of the air.

CHAPTER IV.

THERMODYNAMICS.

169. First Law of Thermodynamics.—The first law of thermodynamics may be thus stated: When heat is transformed into work, or work into heat, the quantity of work is equivalent to the quantity of heat. The experiments of Joule and Rowland establishing this law, and determining the mechanical equivalent, have already been described (§ 114).

170. Second Law of Thermodynamics.—When heat is converted into work by any heat-engine under the conditions that exist on the earth's surface, only a comparatively small proportion of the heat drawn from the source can be so transformed. The remainder is given up to a refrigerator, which in some form must be an adjunct of every heat-engine, and still exists as heat. It will be shown that the heat which is converted into work bears to that which must be drawn from the source of heat a certain simple ratio depending upon the temperatures of the source and refrigerator. The second law of thermodynamics asserts this relation. The ratio between the heat converted into work and that drawn from the source is called the *efficiency* of the engine.

To convert heat into mechanical work, it is necessary that the heat should act through some substance called the *working substance;* as for instance, steam in the steam-engine or air in the hot-air engine. In studying the transformation of heat into work, it is an essential condition that the working substance must, after passing through a cycle of operations, return to the same condition as at the beginning; for if the substance be not in the same condition at the end as at the beginning, internal work may have been done, or internal energy expended, which would increase or diminish the work apparently developed from the heat.

To develop the second law of thermodynamics, we make use of a conception due to Carnot, of an engine completely reversible in all its mechanical and physical operations. In the discussion of the *reversible engine* we employ a principle, first enunciated by Clausius. Clausius' principle is, that heat cannot pass of itself from a cold to a hot body. In many cases this principle agrees with common experience, and in other cases results in accordance with it have been obtained by experiment. It is so fundamental that it is often called the second law of thermodynamics.

Suppose a heat-engine in operation, running forward. It will receive from a source a certain quantity of heat H, transfer to a refrigerator a certain quantity of heat h, and perform a certain amount W of mechanical work. If it be perfectly reversible, it will, by the performance upon it of the amount of work W, take from the refrigerator the quantity of heat h, and restore to the source the amount H. Such an engine will convert into work, under given conditions, as large as possible a proportion of the heat taken from the source. For, let there be two engines, A and B, of which B is reversible, working between the same source and refrigerator. If possible let A perform more work than B, while taking from the source the same amount of heat. If W be the work it performs, and w the work B performs, B will, from its reversibility, by the performance upon it of the work w, less than W, restore to the source the amount of heat, H, which it takes away when running forward. Let A be employed to run B backward: A will take from the source a quantity of heat, H, and perform work, W. B will restore the heat H to the source by the performance upon it of work, w. The system will then continue running, developing the work $W - w$, while the source loses no heat. It must be, then, that A gives up to the refrigerator less heat than B takes away; and the refrigerator must be growing colder. For the purposes of this discussion, we may assume that all surrounding bodies, except the refrigerator, are at the same temperature as the source; hence the work $W-w$, performed by the system of two engines, must be performed by means of heat taken from a body colder than all surrounding bodies. Now this is contrary to the principle of Clausius. The hypothesis with which we started must, therefore, be false; and we must admit that no engine, no contrivance for converting heat into work, can under similar conditions, and while taking the same heat from the source, perform more work than a reversible engine. It follows that all reversible engines, whatever the working substance, have the same efficiency. This is a most important conclusion. In view of it, we may, in studying the conversion of heat into work, choose for the working substance the one which presents the greatest advantage for the study. Since of all substances the properties of gases are best known, we will assume a perfect gas as the working substance. The cycle of four operations which we will study is perfectly reversible. It is known as *Carnot's Cycle*.

Suppose the gas to be enclosed in a cylinder having a tightly-fitting piston. Suppose the cycle to begin by a depression of the piston, compressing the gas, without loss or gain of heat, until the temperature rises from θ to t; where t represents the temperature of the source, and θ that of the refrigerator. In Fig. 53, let Oa represent the volume, and Aa the pressure at the beginning. If the gas be compressed until its volume becomes Ob, its pressure will be bB. AB representing the pressures and corresponding volumes during the operation, is an adiabatic line. This is the *first operation*. For the *second operation*, let the piston rise, and the volume increase

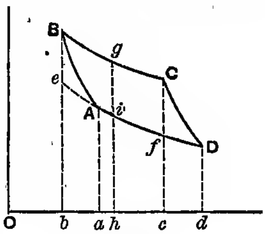

FIG. 53.

from b to c at the constant temperature of the source. The pressure will fall from bB to cC. BC is the isothermal line for the temperature t. During this operation, a quantity of heat represented by H must be taken from the source, to maintain the constant temperature t. For the *third operation* let the piston still ascend, and the volume increase from Oc to Od without loss or gain of heat until the temperature falls from t to θ, the temperature at which the cycle began. CD is an adiabatic line. For the *fourth operation*, let the piston be depressed to the starting point, and the gas maintained at the constant temperature θ of the refrigerator. The volume becomes Oa and the pressure aA, as at the beginning. DA is the isothermal line for the temperature θ.

Now let us consider the work done in each operation. While the piston is being depressed through the volume represented by ab, work must be performed upon it equal to $ab \times$ the mean pressure exerted upon the piston. This mean pressure lies between Aa and Bb, and the product of this by ab is evidently the area $ABba$. In the same way it is shown that when the gas expands from b to c it performs work represented by the area $BCcb$; and again, in the third operation, it performs work represented by $CDdc$. In the fourth operation, when the gas is compressed, work must be done upon it represented by the area $ADda$. During the cycle, therefore, work is done by the gas represented by the area $BCDdb$, and work is done upon the gas represented by the area $BADdb$. The difference represented by the area $ABCD$ is the work done by the engine during the cycle. Since the gas is in all respects in the same condition at the end as at the beginning of the cycle, no work can have been developed from it; and the work which the engine has done must have been derived from the heat communicated to the gas during the second operation.

Now it has been shown that when a gas expands no internal work is done in separating the molecules, and when it expands at constant temperature no change occurs in the internal kinetic energy; the heat which is imparted to the gas during the second operation is, therefore, the equivalent to the work done by the gas upon the piston, and may be represented by the area *BCcb*. It will be seen, also, that the heat given up to the refrigerator during the fourth operation is represented by the area *ADda*, and that heat, the equivalent of the work performed by the engine, represented by the area *ABCD*, has disappeared. Of the heat withdrawn from the source, then, only the fraction $\frac{\text{area } ABCD}{\text{area } BCcb}$ is converted into work. This fraction is the efficiency of the engine.

Now let the operation of the cycle be reversed. Starting with the volume *Oa* the gas expands at the temperature θ, absorbs a quantity of heat represented by h, the same as it gave up when compressed, and performs work represented by *ADda*; next, it is compressed, without loss of heat, until its temperature rises to t, and work represented by *DCcd* is done upon it; next, it is still further compressed at the temperature t, until its volume becomes *Ob*, and its pressure *Bb*. During this operation it gives up the heat H which it absorbed during the direct action, and work represented by *CBbc* is done upon it. Lastly, it expands to the starting-point, and falls to its initial temperature. It will be seen that each operation is the reverse in all respects of the corresponding operation of the direct action, and that during the cycle work represented by the area *ABCD* must be performed upon the engine while the quantity of heat h is taken from the refrigerator, and the quantity of heat H is transferred to the source. Such an engine is therefore a reversible engine; and it converts into work as large a proportion of the heat derived from the source as is possible under the circumstances. An inspection of the figure shows that, since the line *BC* remains the same so long as the amount of heat H and the temperature t of the source remain constant, the only way to increase the proportion of work derived from a given amount of heat H is to increase the difference of temperature between the source and the refrigerator; that is, to increase the area $ABCD$, the line AD must be taken lower down. The proportion of heat which can be converted into work depends, therefore, upon the difference of temperature between source and refrigerator. To determine the nature of this dependence, suppose the range of temperature so small that the sides of the figure $ABCD$ may be considered straight and parallel. Produce AD to e, and draw gh representing the mean pressure for the second operation. Now $ABCD = eBCf = Be \times bc = gi \times bc$. Also $BCcb = gh \times bc$. Then we have

$$\frac{H - h}{H} = \frac{\text{area } ABCD}{\text{area } BCcb} = \frac{gi \times bc}{gh \times bc} = \frac{gi}{gh}.$$

But gh is the pressure corresponding to volume Oh and temperature t, hi is the pressure corresponding to the same volume and temperature θ. These pressures are proportional to the absolute temperatures (§ 156); that is, if t and θ are temperatures on the absolute scale,

$$\frac{ih}{gh} = \frac{\theta}{t},$$

and

$$\frac{gi}{gh} = \frac{H - h}{H} = \frac{t - \theta}{t}: \tag{66}$$

hence

$$\text{efficiency} = \frac{\text{area } ABCD}{\text{area } BCcb} = \frac{t - \theta}{t}. \tag{67}$$

In another form the result contained in Eq. (66) may be written

$$\frac{h}{H} = \frac{\theta}{t}. \tag{68}$$

This proportion has been derived upon the supposition that the range of temperature was very small: but it is equally true for any range; for, let there be a series of engines of small range, of which the second has for a source the refrigerator of the first, the third has for a source the refrigerator of the second, and so on. The first takes from the source the heat H, and gives to the refrigerator the heat h, working between the temperatures t and θ. The second takes the heat h from the refrigerator of the first, and gives to its own refrigerator the heat h_1, working between the temperatures θ and θ_1. The operation of the others is similar; then, from Eq. 68, we have

$$\frac{h}{H} = \frac{\theta}{t},$$

$$\frac{h_1}{h} = \frac{\theta_1}{\theta},$$

$$\frac{h_2}{h_1} = \frac{\theta_2}{\theta_1},$$

$$\frac{h_n}{h_{n-1}} = \frac{\theta_n}{\theta_{n-1}};$$

multiplying, we obtain

$$\frac{h}{H} \times \frac{h_1}{h} \times \frac{h_2}{h_1} \times \dots \frac{h_n}{h_{n-1}} = \frac{\theta}{t} \times \frac{\theta_1}{\theta} \times \frac{\theta_2}{\theta_1} \times \dots \frac{\theta_n}{\theta_{n-1}},$$

or

$$\frac{h_n}{H} = \frac{\theta_n}{t}$$

and

$$\frac{H - h_n}{H} = \frac{t - \theta_n}{t}.$$

Hence it appears that, in a perfect heat-engine, the heat converted into work is to the heat received as the difference of temperature between the source and the refrigerator is to the absolute temperature of the source. This ratio can become unity only when $\theta - 0°$, or when the refrigerator is at the absolute zero of temperature. Since the difference of temperatures between which it is practicable to work is always small compared to the absolute temperature of the source, a perfect heat-engine can convert into work only a small fraction of the heat it receives.

The formulas developed in this section embody what we have called the second law of thermodynamics.

171. Absolute Scale of Temperatures.—An *absolute scale of temperatures*, formed upon the assumed properties of a perfect gas, has already been described (§ 156). No such sub-

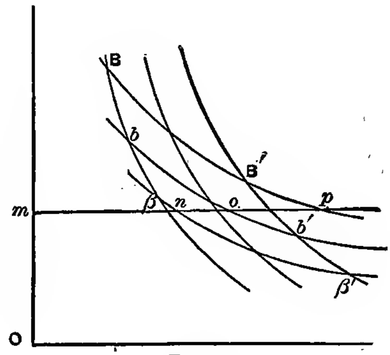

FIG. 54.

stance as a perfect gas exists; but, since (§ 170) any two temperatures on the absolute scale are to each other as the heat taken from the source is to the heat transferred to the refrigerator by a reversible engine, any substance of which we know the properties with sufficient exactness to draw its isothermal and adiabatic lines, may be used as a thermometric substance, and, by means of it, an absolute scale of temperatures may be constructed. For example, in Fig. 54 let BB' be an isothermal line for some substance, corresponding to the temperature t of boiling water at a standard pressure. Let $\beta\beta'$ be the isothermal line for the temperature t_0 of melting ice, and let bb' be an isothermal line for an intermediate temperature. Let $B\beta$, $B'\beta'$, be adiabatic lines, such that, if the substance expand at constant temperature t from the condition B to the condition B', the equivalent in heat of one mechanical unit of energy will be absorbed. Now, the figure $BB'\beta'\beta$ represents Carnot's cycle, and the heat given to the refrigerator at the temperature t_0, measured in mechanical units, is less than the heat taken from the source at the temperature t, by the energy represented by the area $BB'\beta'\beta$; or, the heat given to the refrigerator is equal to 1 — area $B\beta'$: hence

$$\frac{t_0}{t} = \frac{1 - area\ B\beta'}{1},$$

and

$$\frac{t - t_0}{t} = \frac{area\ B\beta'}{1}.$$

Now, if $t - t_0 = 100°$, as in the Centigrade scale, we have

$$\frac{100}{t} = \frac{area\ B\beta'}{1},$$

and

$$t = \frac{100}{area\ B\beta'}. \tag{69}$$

If θ be the temperature corresponding to the isothermal line bb', we have, as above,

$$\frac{\theta}{t} = \frac{1 - area\ Bb'}{1},$$

whence

$$\theta = t(1 - area\ Bb') = \frac{100}{area\ B\beta'}(1 - area\ Bb'). \qquad (70)$$

If, now, it be proposed to use the substance as a thermometric substance by noting its expansion at constant pressure, take Om to represent that pressure, and draw the horizontal line $mnop$; mn is the volume of the substance at temperature t_0, mo the volume at temperature θ, and mp the volume at temperature t.

This method of constructing an absolute scale of temperature was proposed by Thomson.

172. The Steam-Engine.—The steam-engine in its usual form consists essentially of a piston, moving in a closed cylinder, which is provided with passages and valves by which steam can be admitted and allowed to escape. A boiler heated by a suitable furnace supplies the steam. The valves of the cylinder are opened and closed automatically, admitting and discharging the steam at the proper times to impart to the piston a reciprocating motion, which may be converted into a circular motion by means of suitable mechanism.

There are two classes of steam-engines, condensing and non-condensing. In condensing engines the steam, after doing its work in the cylinder, escapes into a condenser, kept cold by a circulation of cold water. Here the steam is condensed into water; and this water, with air or other contents of the condenser, is removed by an "air-pump." In non-condensing engines the steam escapes into the open air. In this case the temperature of the refrigerator must be considered at least as high as that of saturated steam at the atmospheric pressure, or about 100°, and the temperature of the source must be taken as that of saturated steam at the boiler pressure. Applying the expression for the efficiency (§ 170),

$$e = \frac{t - \theta}{t},$$

it will be seen, that, for any boiler pressure which it is safe to employ in practice, it is not possible, even with a perfect engine, to convert into work more than about fifteen per cent of the heat used.

In the condensing engine the temperature of the refrigerator may be taken as that of saturated steam at the pressure which exists in the condenser, which is usually about 30° or 40°: hence $t - \theta$ is a much larger quantity for condensing than for non-condensing engines. The gain of efficiency is not, however, so great as would appear from the formula, because of the energy that must be expended to maintain the vacuum in the condenser.

173. **Hot-air and Gas Engines.**—Hot-air engines consist essentially of two cylinders of different capacities, with some arrangement for heating air in, or on its way to, the larger cylinder. In one form of the engine, an air-tight furnace forms the passage between the two cylinders, of which the smaller may be considered as a supply-pump for taking air from outside, and forcing it through the furnace into the larger cylinder, where, in consequence of its expansion by the heat, it is enabled to perform work. On the return stroke, this air is expelled into the external air, still hot, but at a lower temperature than it would have been had it not expanded and performed work. This case is exactly analogous to that of the steam-engine, in which water is forced by a piston working in a small cylinder,

into a boiler, is there converted into steam, and then, acting upon a much larger piston, performs work, and is rejected. In another form of the engine, known as the "ready motor," the air is forced into the large cylinder through a passage kept supplied with crude petroleum. The air becomes saturated with the vapor, forming a combustible mixture, which is burned in the cylinder itself.

The Stirling hot-air engine and the Rider "compression engine" are interesting as realizing an approach to Carnot's cycle.

These engines, like those described above, consist of two cylinders of different capacities, in which work air-tight pistons; but, unlike those, there are no valves communicating with the external atmosphere. Air is not taken in and rejected; but the same mass of air is alternately heated and cooled, alternately expands and contracts, moving the piston, and performing work at the expense of a portion of the heat imparted to it.

It is of interest to study a little more in detail the cycle of operations in these two forms of engines. The larger of the two cylinders is kept constantly at a high temperature by means of a furnace, while the smaller is kept cold by the circulation of water. The cylinders communicate freely with each other. The pistons are connected to cranks set on an axis, so as to make an angle of nearly ninety degrees with each other. Thus both pistons are moving for a short time in the same direction twice during the revolution of the axis. At the instant that the small piston reaches the top of its stroke, the large piston will be near the bottom of the cylinder, and descending. The small piston now descends, as well as the large one, the air in both cylinders is compressed, and there is but little transfer from one to the other. There is, therefore, comparatively little heat given up. The large piston, reaching its lowest point, begins to ascend, while the descent of the smaller continues. The air is rapidly transferred to the larger heated cylinder, and expands while taking heat from the highly heated surface. After the small piston has reached its lowest point, there is a short time during which both the pistons are rising and the air expanding, with but little transfer from one cylinder to the other, and with a relatively small absorption of heat. When the descent of the large piston begins, the small one still rising, the air is rapidly transferred to the smaller cylinder: its volume is diminished, and its heat is given up to the cold surface with which it is brought in contact. The completion of this operation brings the air back to the condition from which it started. It will be seen that there are here four operations, which, while not presenting the simplicity of the four operations of Carnot,—since the first and third are not performed without transfer of heat, and the second and fourth not without change of temperature,—still furnish an example of work done by heat through a series of changes in the working substance, which brings it back, at the end of each revolution, to the same condition as at the beginning.

Gas-engines derive their power from the force developed by the combustion, within the cylinder, of a mixture of illuminating gas and air.

As compared with steam-engines, hot-air and gas engines use the working substance at a much higher temperature. $t - \theta$ is, therefore, greater, and the theoretical efficiency higher. There are, however, practical difficulties connected with the lubrication of the sliding surfaces at such high temperatures that have so far prevented the use of large engines of this class.

174. Sources of Terrestrial Energy.—Water flowing from a higher to a lower level furnishes energy for driving machinery. The energy theoretically available in a given time is the weight of the water that flows during that time multiplied by the height of the fall. If this energy be not utilized, it develops heat by friction of the water or of the material that may be transported by it. But water-power is only possible so long as the supply of water continues. The supply of water is dependent upon the rains; the rains depend upon evaporation; and evaporation is maintained by solar heat. The energy of *water-power* is, therefore, transformed solar energy.

A moving mass of air possesses energy equal to the mass multiplied by half the square of the velocity. This energy is available for propelling ships, for turning windmills, and for other work. Winds are due to a disturbance of atmospheric equilibrium by solar heat; and the energy of *wind-power*, like that of water-power, is, therefore, derived from solar energy.

The *ocean currents* also possess energy due to their motion, and this motion is, like that of the winds, derived from solar energy.

By far the larger part of the energy employed by man for his purposes is derived from the combustion of wood and coal. This energy exists as the potential *energy of chemical separation* of oxygen from carbon and hydrogen. Now, we know that vegetable matter is formed by the action of the solar rays through the mechanism of the leaf, and that coal is the carbon of plants that grew and decayed in a past geological age. The energy of wood and coal is, therefore, the transformed energy of solar radiations.

It is well known that, in the animal tissues, a chemical action takes place similar to that involved in combustion. The oxygen taken into the lungs and absorbed by the blood combines by processes with which we are not here concerned with the constituents of the food. Among the products of this combination are carbon dioxide and water, as in the combustion of the same substances elsewhere. Lavoisier assumed that such chemical combinations were the source of *animal heat*, and was the first to attempt a measurement of it. He compared the heat developed with that due to the formation of the carbonic dioxide exhaled in a given time. Despretz and Dulong made similar experiments with more perfect apparatus, and found that the heat produced by the animal was about one tenth greater than would have been produced by the formation by combustion of the carbonic acid and water exhaled.

These and similar experiments, although not taking into account all the chemical actions taking place in the body, leave no doubt that animal heat is due to atomic and molecular changes within the body.

The work performed by muscular action is also the transformed energy of food. Rumford, in 1798, saw this clearly; and he showed, in a paper of that date, that the amount of work done by a horse is much greater than would be obtained by using its food as fuel for a steam-engine.

Mayer, in 1845, held that an animal is a heat-engine, that every motion of the animal is a transformation into work of the heat developed in the tissues.

Hirn, in 1858, executed a series of interesting experiments bearing upon this subject. In a closed box was placed a sort of treadmill, which a man could cause to revolve by stepping from step to step. He thus performed work which could be measured by suitable apparatus outside the box. The treadwheel could also be made to revolve backward by a motor placed outside; when the man descended from step to step, and work was performed upon him.

Three distinct experiments were performed; and the amount of oxygen consumed by respiration, and the heat developed, were determined.

In the first experiment the man remained in repose; in the second he performed work by causing the wheel to revolve; in the third the wheel was made to revolve backward, and work was performed upon him. In the second experiment, the amount of heat developed for a gram of oxygen consumed was much less, and in the third case much greater, than in the first; that is, in the first case, the heat developed was due to a chemical action, indicated by the absorption of oxygen; in the second, a portion of the chemical action went to perform the work, and hence a less amount of heat was developed; while in the third case the motor, causing the treadwheel to revolve, performed work, which produced heat in addition to that due to the chemical action.

It has been thought that muscular energy is due to the waste of the muscles themselves: but experiments show that the waste of nitrogenized material is far too small in amount to account for the energy developed by the animal; and we must, therefore, conclude that the principal source of muscular energy is the oxidation of the non-nitrogenized material of the blood by the oxygen absorbed in respiration.

An animal is, then, a machine for converting the potential energy of food into mechanical work: but he is not, as Mayer supposed, a heat-engine; for he performs far more work than could be performed by a perfect heat-engine, working between the same limits of temperature, and using the food as fuel.

The food of animals is of vegetable origin, and owes its energy to the solar rays. Animal heat and energy is, therefore, the transformed energy of the sun.

The *tides* are mainly caused by the attraction of the moon upon the waters of the earth. If the earth did not revolve upon its axis, or, rather, if it always presented one face to the moon, the elevated waters would remain stationary upon its surface, and furnish no source of energy. But as the earth revolves, the crest of the tidal wave moves apparently in the opposite direction, meets the shores of the continents, and forces the water up the bays and rivers, where energy is wasted in friction upon the shores or may be made use of for turning mill-wheels. It is evident that all the energy derived from the tides comes from the rotation of the earth upon its axis; and a part of the energy of the earth's rotation is, therefore, being dissipated in the heat of friction they cause.

The internal heat of the earth and a few other forms of energy, such as that of native sulphur, iron, etc., are of little consequence as sources of useful energy. They may be considered as the remnants of the original energy of the earth.

175. Energy of the Sun.—It has been seen that the sun's rays are the source of all the forms of energy practically available, except that of the tides. It has been estimated that the heat received by the earth from the sun each year would melt a layer of ice over the entire globe a hundred feet in thickness. This represents energy equal to one horse-power for each fifty square feet of surface, and the heat which reaches the earth is only one twenty-two-hundred-millionth of the heat that leaves the sun. Notwithstanding this enormous expenditure of energy, Helmholtz and Thomson have shown that the nebular hypothesis, which supposes the solar system to have originally existed as a chaotic mass of widely separated gravitating particles, presents to us an adequate source for all the energy of the system. As the particles of the system rush together by their mutual attractions, heat is generated by their collision; and after they have collected into large masses, the condensation of these masses continues to generate heat.

176. Dissipation of Energy.—It has been seen that only a fraction of the energy of heat is available for transformation into other forms of energy, and that such transformation is possible only when a difference of temperature exists. Every conversion of other forms of energy into heat puts it in a form from which it can be only partially recovered. Every transfer of heat from one body to another, or from one part to another of the same body, tends to equalize temperatures, and to diminish the proportion of energy available for transformation. Such transfers of heat are continually taking place; and, so far as our present knowledge goes, there is a tendency toward an equality of temperature, or, in other words, a uniform molecular motion, throughout the universe. If this condition of things were reached, although the total amount of energy existing in the universe would remain unchanged, the possibility of transformation would be at an end, and all activity and change would cease. This is the doctrine of the dissipation of energy to which our limited knowledge of the operations of nature leads us; but it must be remembered that our knowledge is very limited, and that there may be in nature the means of restoring the differences upon which all activity depends.

MAGNETISM AND ELECTRICITY.

CHAPTER I.

MAGNETISM.

177. Fundamental Facts.—Masses of iron ore are sometimes found which possess the property of attracting pieces of iron and a few other substances. Such masses are called *natural magnets* or lodestones. A bar of steel may be so treated as to acquire similar properties. It is then called a *magnet*. Such a magnetized steel bar may be used as fundamental in the investigation of the properties of magnetism.

If pieces of iron or steel be brought near a steel magnet, they are attracted by it, and unless removed by an outside force they remain permanently in contact with it. While in contact with the magnet, the pieces of iron or steel also exhibit magnetic properties. The iron almost wholly loses these properties when removed from the magnet. The steel retains them and itself becomes a magnet. The reason for this difference is not known. It is usually said to be due to a *coercive force* in the steel. The attractive power of the original magnet for other iron or steel remains unimpaired by the formation of new magnets.

A body which is thus magnetized or which has its magnetic condition disturbed is said to be affected by ***magnetic induction***.

In an ordinary bar magnet there are two small regions, near the ends of the bar, at which the attractive powers of the magnet are most strongly manifested. These regions are called the *poles* of the magnet. The line joining two points in these regions, the location of which will hereafter be more closely defined, is called the *magnetic axis*. An imaginary plane drawn normal to the axis at its middle point is called the *equatorial plane*.

If the magnet be balanced so as to turn freely in a horizontal plane, the axis assumes a direction which is approximately north and south. The pole toward the north is usually called the north or positive pole; that toward the south, the south or negative pole.

If two magnets be brought near together, it is found that their like poles repel and unlike poles attract one another.

If the two poles of a magnet be successively placed at the same distance from a pole of another magnet, it is found that the forces exerted are equal in amount and oppositely directed.

The direction assumed by a freely suspended magnet shows that the earth acts as a magnet, and that its north magnetic pole is situated in the southern hemisphere.

If a bar magnet be broken, it is found that two new poles are formed, one on each side of the fracture, so that the two portions are each perfect magnets. This process of making new magnets by subdivision of the original one may be, so far as known, continued until the magnet is divided into its least parts, each of which will be a perfect magnet.

This last experiment enables us at once to adopt the view that the properties of a magnet are due to the resultant action of its constituent magnetic molecules.

178. Law of Magnetic Force.—By the help of the torsion balance, the principle of which is described in §§ 82, 188, and by using very long, thin, and uniformly magnetized bars, in which the poles can be considered as situated at the extremities, Coulomb showed that the repulsion between two similar poles, and the attraction between two dissimilar poles, is inversely as the square of the distance between them.

Coulomb also demonstrated the same law by another method. He suspended a short magnet so that it could oscillate about its centre in the horizontal plane. He first observed the time of its oscillation when it was oscillating in the earth's magnetic field. He then placed a long magnet vertically, so that one of its poles was in the horizontal plane of the suspended magnet, and in the magnetic meridian passing through its centre, and observed the times of oscillation when the pole of the vertical magnet was at two different distances from the suspended magnet. If we represent by I the moment of inertia of the suspended magnet, by M its magnetic moment, by H the horizontal intensity of the earth's magnetism, by h_1 and h_2 the force in the region occupied by the suspended magnet due to the vertical magnet in its two positions, it may be shown as in § 183 that the times of oscillation of the suspended magnet should be respectively $t_H = \pi \sqrt{\frac{I}{HM}}$, $t_{h_1} = \pi \sqrt{\frac{I}{h_1 H}}$, $t_{h_2} = \pi \sqrt{\frac{I}{h_2 M}}$. From such equations, by elimination of H, the values of h_1 and h_2 were obtained, and were found to be in accordance with the law of magnetic force already given.

All theories of magnetism assume that the force between two magnet poles is proportional to the product of the strengths of the poles. The law of magnetic force is then the same as that upon which the discussion of potential (§§ 28, 29) was based. The theorems there discussed are in general applicable in the study of magnetism, although modifications in the details of their application occur, arising from the fact that the field of force about a magnet is due to the combined action of two dissimilar and equal poles.

If m and m_1 represent the strengths of two magnet poles, r

the distance between them, and k a factor depending on the units in which the strength of the pole is measured, the formula expressing the force between the poles is $k\frac{mm_{\prime}}{r^2}$.

179. Definitions of Magnetic Quantities.—The law of magnetic force enables us to define a *unit magnet pole*, based upon the fundamental mechanical units.

If two perfectly similar magnets, infinitely thin, uniformly and longitudinally magnetized, be so placed that their positive poles are unit distance apart, and if these poles repel one another with unit force, the magnet poles are said to be of *unit strength*. Hence, in the expression for the force between two poles, k becomes unity, and the dimensions of $\frac{m^2}{r^2}$ are those of a force. That is,

$$\left[\frac{m^2}{r^2}\right] = MLT^{-2},$$

from which the dimensions of a magnet pole are

$$[m] = M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-1}.$$

This definition of a unit magnet pole is the foundation of the *magnetic system of units*. The *strength* of a magnet pole is then equal to the force which it will exert on a unit pole at unit distance.

The product of the strength of the positive pole of a uniformly and longitudinally magnetized magnet into the distance between its poles is called its *magnetic moment*.

The quotient of the magnetic moment of such a magnet by its volume, or the magnetic moment of unit of volume, is called the *intensity of magnetization*.

The dimensions of magnetic moment and of intensity of magnetization follow from these definitions. They are respectively

$$[ml] = M^{\frac{1}{2}}L^{\frac{5}{2}}T^{-1} \quad \text{and} \quad \left[\frac{ml}{l^3}\right] = M^{\frac{1}{2}}L^{-\frac{1}{2}}T^{-1}.$$

180. Distribution of Magnetism in a Magnet.—If we conceive of a single row of magnetic molecules with their unlike poles in contact, we can easily see that all the poles, except those at the ends, neutralize one another's action, and that such a row will have a free north pole at one end and a free south pole at the other. If a magnet be thought of as made up of a combination of such rows of different lengths, the action of their free poles may be seen to be the same as that of an imaginary distribution of equal quantities of north and south *magnetism* on the surface and throughout the volume of the magnet. If the magnet be uniformly magnetized, the volume distribution becomes zero. The *surface distribution* of magnetism will sometimes be used to express the magnetization of a magnet. In that case what has hitherto been called the magnetic intensity becomes the *magnetic density*. It is defined as the ratio of the quantity of magnetism on an element of surface to the area of that element. To illustrate this statement, we will consider an infinitely thin and uniformly magnetized bar, of which the length and cross-section are represented by l and s respectively. Its magnetic intensity is $\frac{ml}{ls}$ or $\frac{m}{s}$. If, now, for the pole m we substitute a continuous surface distribution over the end of the bar, then $\frac{m}{s}$ is also the density of that distribution.

The dimensions of magnetic density follow from this definition. They are

$$\left[\frac{m}{s}\right] = \frac{M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-1}}{L^2} = M^{\frac{1}{2}}L^{-\frac{1}{2}}T^{-1}.$$

Coulomb showed, by a method of oscillations similar to that described in § 178, that the magnetic force at different points along a straight bar magnet gradually increases from the middle of the bar, where it is imperceptible, to the extremities. This would not be the case if the bar magnet were made up of equal straight rows of magnetic molecules in contact, placed side by side. With such an arrangement there would be no force at any point along the bar, but it would all appear at the two ends. The mutual interaction of the molecules of contiguous rows make such an arrangement, however, impossible.

In the earth's magnetic field, in which the lines of magnetic force may be considered parallel, a couple will be set up on any magnet, so magnetized as to have only two poles, due to the action of equal quantities of north and south magnetism distributed in the magnet. The points at which the forces making up this couple are applied are the *poles* of the magnet, and the line joining them is the *magnetic axis*. These definitions are more precise than those which could be given at the outset.

181. Action of One Magnet on Another.—The investigation of the mechanical action of one magnet on another is important in the construction of apparatus for the measurement of magnetism.

(1) To determine the *potential of a short bar magnet* at a point distant from it, let NS (Fig. 55) represent the magnet of length $2l$, the poles of which are of strength m, and let the point P be at a distance r from the centre of the magnet, taken as origin. Let the x axis coincide with the axis of the magnet. The potential at P is then

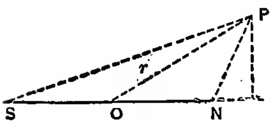

Fig. 55.

$$V = m\left(\frac{1}{(y^2 + (x - l)^2)^{\frac{1}{2}}} - \frac{1}{(y^2 + (x + l)^2)^{\frac{1}{2}}}\right)$$

$$= m\left(\frac{1}{(r^2 + l^2 - 2xl)^{\frac{1}{2}}} - \frac{1}{(r^2 + l^2 + 2xl)^{\frac{1}{2}}}\right).$$

This expression expanded gives

$$V=\frac{2mlx}{r^3}-\frac{3ml^3x}{r^5}+\frac{5ml^3x^3}{r^7}, \tag{71}$$

if we assume r so large that we may neglect terms of higher order in l. The first term is the most important, and if r be very great compared with l, the other terms may be neglected. The ratio $\frac{x}{r}$ is the cosine of the angle PON or θ. If we represent the magnetic moment $2ml$, as is generally done, by M, the potential at any very distant point becomes

$$\frac{M}{r^2}\cos\theta. \tag{72}$$

Since $\cos\theta$ is zero for all points in a plane through the origin at right angles to the magnetic axis, that plane is an equipotential surface of zero potential. It is the plane defined as the equatorial plane. The lines of force evidently originate at the poles and pass perpendicularly through this surface. This system of lines of force can be easily illustrated by scattering fine iron filings on a sheet of paper held over a bar magnet. They will arrange themselves approximately along the lines of force.

At a point on the line of the axis where $r=x$, the potential becomes

$$V=\frac{M}{x^2}+\frac{Ml^2}{x^4}+\ldots \tag{73}$$

(2) In one method of application of the instrument called the magnetometer it is necessary to know the expression for the *moment of couple* set up by the action of a magnet at right

angles to another, the centre of which is in the prolongation of the axis of the first magnet. Let the centre of the first magnet be the origin, and its axis the x axis. Represent the strength of its poles by m, and the strength of the pole of the second magnet by $m_{\prime}$, the lengths of the two magnets by $2l$ and $2y$ respectively. To determine the moment of couple due to the action of the first magnet on the second, we must first find the component along the x axis of the force due to the first magnet on a pole $m_{\prime}$, at a point distant y from the x axis.

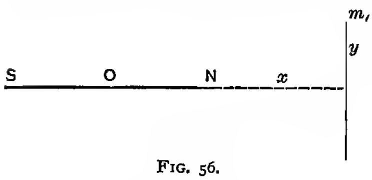

FIG. 56.

The force due to the pole of the first magnet at N (Fig. 56) on a pole $m_{\prime}$ is

$$\frac{mm_{\prime}}{y^2 + (x - l)^2}.$$

The cosine of the angle made by this force with the x axis is $\dfrac{x - l}{(y^2 + (x - l)^2)^{\frac{1}{2}}}$. Hence the component of this force along the x axis is

$$\frac{mm_{\prime}(x - l)}{(y^2 + (x - l)^2)^{\frac{3}{2}}}.$$

Hence the component along the x axis of the whole force on the pole $m_{\prime}$, due to the first magnet, is

$$mm_{\prime}\left(\frac{x - l}{(y^2 + (x - l)^2)^{\frac{3}{2}}} - \frac{x + l}{(y^2 + (x + l)^2)^{\frac{3}{2}}}\right).$$

When this expression is expanded in increasing negative powers of x, neglecting all terms containing higher powers of x than the fifth, we obtain

$$4mm_{\prime}l\left(\frac{1}{x^3} + \frac{2l^2}{x^5} - \frac{3y^2}{x^5}\right).$$

An equal and oppositely directed component acts upon the other pole $-m_{,}$ of the second magnet. Hence the moment of couple due to the action of the first magnet upon the second is

$$8mm_{,}ly\left(\frac{1}{x^3}+\frac{2l^2}{x^5}-\frac{3y^2}{x^6}\right). \qquad (74)$$

If y be such that $3y^2 = 2l^2$, or if the ratio of the lengths of the two magnets used be $1 : \sqrt{1.5}$, the second and third terms vanish, and the expression for the moment of couple depends only on the first term of the series. In practice it is not possible to completely neglect the other terms, on account of the uncertainty as to the position of the poles in the figure of a magnet, but by making the lengths of the two magnets as 1 to $\sqrt{1.5}$, the numerator of the term having x^5 in the denominator is made very small, and is eliminated by the method of observation employed, as will be explained in the discussion of the magnetometer.

182. The Magnetic Shell.—A *magnetic shell* may be defined as an infinitely thin sheet of magnetizable matter, magnetized transversely; so that any line in the shell normal to its surfaces may be looked on as an infinitesimally short and thin magnet. These imaginary magnets have their like poles contiguous. The product of the intensity of magnetization at any point in the shell into the thickness of the shell at that point is called the ***strength of the shell*** at that point, and is denoted by the symbol j.

The dimensions of the strength of a magnetic shell follow at once from this definition. We have $[j]$ equal to the dimensions of intensity of magnetization multiplied by a length. Therefore $[j] = M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-1}$.

We obtain first the potential of such a shell of infinitesimal area. Let the origin (Fig. 57) be taken half-way between the two faces of the shell, and let the shell stand perpendicular to the x axis. Let a represent the area of the shell, supposed infinitesimal, $2l$ the thickness of the shell, and d the magnetic intensity. The volume of this infinitesimal magnet is $2al$, and from the definition of magnetic intensity $2ald$ is its magnetic moment. The potential at the point P is then given by Eq. 72, since l is so small that all but the first term in the series of Eq. 71 may be neglected. We have

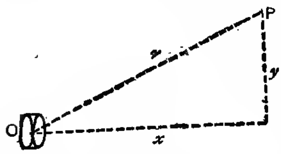

FIG. 57.

$$V = \frac{M}{r^2}\cos\theta = \frac{2ald}{r^2}\cos\theta.$$

Now $a\cos\theta$ is the projection of the area of the shell upon a plane through the origin normal to the radius vector r, and, since a is infinitesmal, $\frac{a\cos\theta}{r^2}$ is the solid angle ω bounded by the lines drawn from P to the boundary of the area a. The potential then becomes $V = 2ald\omega = j\omega$, since $2ald$ is what has been called the strength of the shell.

The same proof may be extended to any number of contiguous areas making up a finite magnetic shell. The potential due to such a shell is then $\Sigma j\omega$. If the shell be of uniform strength, the potential due to it becomes $j\Sigma\omega$ and is got by summing the elementary solid angles. This sum is the solid angle Ω, bounded by the lines drawn from the point of which the potential is required to the boundary of the shell. The potential due to a magnetic shell of uniform strength is therefore

$$j\Omega. \tag{75}$$

It is independent of the form of the shell, and dependent only on the form of its boundary. At a point very near the positive face of a flat shell, so near that the solid angle subtended by the shell equals 2π, the potential is $2\pi j$; at a point in the plane of the shell outside its boundary where the angle subtended is zero, the potential is zero; and near the other or negative face of the shell it is $-2\pi j$. The whole work done, then, in moving a unit magnet pole from a point very near one face to a point very near the other face is $4\pi j$. This result is of importance in connection with electrical currents.

183. Magnetic Measurements.—It was shown by Gilbert in a work published in 1600, that the earth can be considered as a magnet, having its positive pole toward the south and its negative toward the north. The determination of the magnetic relations of the earth are of importance in navigation and geodesy. The principal magnetic elements are the declination, the dip, and the horizontal intensity.

The *declination* is the angle between the magnetic meridian, or the direction assumed by the axis of a magnetic needle suspended to move freely in a horizontal plane, and the geographical meridian.

The *dip* is the angle made with the horizontal by the axis of a magnetic needle suspended so as to turn freely in a vertical plane containing the magnetic meridian.

The *horizontal intensity* is the strength of the earth's magnetic field resolved along the horizontal line in the plane of the magnetic meridian. A magnet pole of strength m in a field in which the horizontal intensity is represented by H is urged along this horizontal line with a force equal to mH. From this equation the dimensions of the horizontal intensity, and so also of the strength of a magnetic field in any case, are

$$[H] = \left[\frac{MLT^{-2}}{m}\right] = M^{\frac{1}{2}}L^{-\frac{1}{2}}T^{-1}.$$

The horizontal intensity can be measured relatively to some assumed magnet as standard, by allowing the magnet to oscillate freely in the horizontal plane about its centre, and noting the time of oscillation. The relation between the magnetic moment M of the magnet and the horizontal intensity H is calculated by a formula analogous to that employed in the computation of g from observations with the pendulum. If the magnet be slightly displaced from its position of equilibrium, so as to make small oscillations about its point of suspension, it can be shown as in § 39 that it is describing a simple harmonic motion, and as in § 41 (1) that the kinetic energy of the magnet when its axis coincides, during an oscillation, with the magnetic meridian is

$$\tfrac{1}{2}I\frac{4\pi^2\phi^2}{T^2}.$$

The potential energy at the extremity of its arc is due to the magnetic force mH acting on the poles. The component of this force which is efficient in moving the magnet is $mH \sin \alpha$ or $mH\alpha$, if α be always very small. Since α varies between 0 and ϕ, the average force sufficient in turning the needle is $\frac{1}{2}mH\phi$. The poles upon which this force acts move from the position of maximum kinetic energy to the position of no kinetic energy, through a distance $l\phi$, if l represent the half length of the magnet. The potential energy of the couple formed by the two poles of the magnet is then $mHl\phi^2$, and this is equal to the kinetic energy at the point of equilibrium; that is,

$$\tfrac{1}{2}I\frac{4\pi^2\phi^2}{T^2} = mHl\phi^2.$$

Hence if we write $2ml = M$, the magnetic moment of the magnet, we obtain $MH = \frac{4\pi^2 I}{T^2}$; or if we take the time of oscillation as $t = \frac{T}{2}$, we have

$$MH = \frac{\pi^2 I}{t^2}. \tag{76}$$

The moment of inertia I may be either computed directly from the magnet itself, if it be of symmetrical form, or it may be determined experimentally by the method of § 36, Eq. 23, which applies in this case. The horizontal intensity is then determined relative to the magnetic moment of the assumed standard magnet.

This measure may be used to give an absolute measure of H by combining with it another observation which gives an independent relation between M and H. In one arrangement of the apparatus two magnets are used: one, the *deflected magnet*, so suspended as to turn freely in the horizontal plane; and the other, the *deflecting magnet*, the one of moment M used in the last operation, carried upon a bar which can be turned about a vertical axis passing through the point of suspension of the deflected magnet. The centre of the deflected magnet is in the prolongation of the axis of the deflecting magnet, and, when the apparatus is used, the carrier bar is turned until the two magnets are at right angles to one another. The equilibrium established is due to two couples acting on the deflected magnet, one arising from the action of the earth's magnetism, and the other from that of the deflecting magnet. This latter has been already discussed in § 181. The couple acting on the deflected magnet is expressed by $4Mm_{,}y\left(\frac{1}{x^3} + \frac{P}{x^5}\right)$, where P represents the small numerator of the correction term. This correction can be made very small in practice by giving to the magnet, as already explained, lengths in the ratio of 1 to $\sqrt{1.5}$. The opposing equal couple is $2m_{,}Hy \sin\phi$, where ϕ represents the angle of deflection from the magnetic meridian. We have then $4Mm_{,}y\left(\frac{1}{x^3}+\frac{P}{x^5}\right) = 2my_{,}H \sin\phi$, or $\frac{1}{x^3}+\frac{P}{x^5} = \frac{1}{2}\frac{H}{M}\sin\phi$. Since P is always a very small quantity, this equation may be written

$$\frac{M}{H} = \tfrac{1}{2}x^3 \sin\phi\left(1 - \frac{P}{x^2}\right). \qquad (77)$$

P is determined by measuring the angles ϕ and $\phi_{,}$ for two different distances x and $x_{,}$. The equations containing the results of these measurements are

$$\frac{M}{H} = \tfrac{1}{2}x^3 \sin\phi\left(1 - \frac{P}{x^2}\right)$$

and

$$\frac{M}{H} = \tfrac{1}{2}x_{,}^3 \sin\phi_{,}\left(1 - \frac{P}{x_{,}^2}\right).$$

From these equations the value of P is found to be equal to

$$\frac{\tfrac{1}{2}x^3 \sin\phi - \tfrac{1}{2}x_{,}^3 \sin\phi_{,}}{\tfrac{1}{2}x \sin\phi - \tfrac{1}{2}x_{,} \sin\phi_{,}}.$$

By substitution of this value of P in either of the above equations, the value of $\frac{M}{H}$ is obtained in absolute units. By combination of Eq. (76) and Eq. (77) the value of H is obtained independent of M, and in absolute units.

It is evident that the value of M can be obtained also in absolute units from the same equations.

In determinations of the horizontal intensity in which great accuracy is desired, corrections must be introduced in these equations for the changes of magnetic moment due to changes of temperature (§ 185) and to induction (§ 184).

184. Magnetic Induction.—In the foregoing discussions the effect of magnetic induction has been neglected, and the magnets considered are those known as permanent magnets. Phenomena, however, arise when bodies not permanently magnetized are brought into a magnetic field, which are due to magnetic induction. It was found by Faraday that all bodies are affected by the presence of a magnet. Some of them, such as iron, nickel, cobalt, and oxygen, seem to be attracted by the magnet. Others, such as bismuth, copper, most organic substances, and nitrogen, seem to be repelled from the magnet. The former are said to be *ferromagnetic* or *paramagnetic;* the latter, *diamagnetic.*

The most obvious explanation of these phenomena, and the one adopted by Faraday, is to ascribe them to a distribution of the induced magnetization in paramagnetic bodies, in an opposite direction from that in diamagnetic bodies. If a paramagnetic body be brought between two opposite magnet poles, a north pole is induced in it near the external south pole, and a south pole near the external north pole. The magnetic separation is then said to be in the direction of the lines of force. According to this explanation, then, the separation of the induced magnetization in a diamagnetic body is in a direction opposite to that of the lines of force. In other words, if a diamagnetic body be brought between two opposite magnet poles, the explanation asserts that a north pole is induced in it near the external north pole, and a south pole near the external south pole.

One of Faraday's experiments, however, indicates that the different behavior of bodies of these two classes may be due only to a more or less intense manifestation of the same action.

He found that a solution of ferrous sulphate, sealed in a glass tube, behaves, immersed in a weaker solution of the same salt, as a paramagnetic body; but, when immersed in a stronger solution, as a diamagnetic body. It may, from this experiment, be concluded that the direction of the induced magnetization is the same for all bodies, and that the exhibition of diamagnetic or paramagnetic properties depends, not upon the direction of induced magnetization, but upon the greater or less intensity of magnetization of the surrounding medium.

Faraday discovered that many bodies, while in a vacuum, exhibit diamagnetic properties. In accordance with this explanation, we must conclude that a vacuum can have magnetic properties. It seemed to Faraday unlikely that this should be the case, and he therefore adopted the explanation which was first given. As it has been since shown that the ether which serves as a medium for the transmission of light, and which pervades every so-called vacuum, is also probably concerned in electrical and magnetic phenomena, there is no longer any reason for the opinion that the possession of magnetic properties by a vacuum is inherently improbable. In accordance with this view, in what follows we shall adopt the second explanation, which was developed by Thomson.

In order to express the difference between paramagnetic and diamagnetic bodies it is necessary to use some definitions which did not appear in the treatment of permanent magnets. To understand these it is necessary to determine the magnetic force within a magnet, as upon it depends the induced magnetization. The force in the interior of a magnet is measured by considering an infinitely short cylinder or thin disk, the axis of which is parallel to the axis of the magnet, cut out of the interior of the magnet. The force exerted upon a magnet pole within this space is the force to be considered. Assume a straight bar magnet, uniformly and longitudinally magnetized, and suppose such a disk cut out within it, with faces perpendicular to the magnetic axis. We may then assume (§ 180) that there will be a uniform distribution of magnetism on both faces of the disk, positive on one face and negative on the other. The force due to this imaginary distribution on a pole in the centre of the disk will be twice that due to one face. If its density be called i, by § 29 (3) the force due to one face is $2\pi i$. If there be besides a magnetic force F in the field, the total force on the pole is $F+4\pi i$. If the straight bar be not originally magnetized, and be placed in a uniform magnetic field, it is assumed that i is proportional to F. Let $i=kF$, and call k the *coefficient of induced magnetization.* We have then the total force within the cavity equal to $(1+4\pi k)F$. This quantity is called by Maxwell the *magnetic induction*, and the factor $1+4\pi k$ the *magnetic inductive capacity* of the substance. Thomson and Rowland call it the magnetic permeability.

Now, to make clear what is meant by the classification of bodies as paramagnetic and diamagnetic, we may proceed as follows. Suppose an infinitesimal cube of the substance to be tested placed in a magnetic field which is not uniform, with one of its faces normal to the lines of force. The intensity of the induced magnetization will be kF, if we assume that the cube is so small that we may neglect the variation of its induced magnetization, due to the variation within it of the magnetic force F. In most bodies k is so small that we may also assume that the induced magnetization does not appreciably alter the field, and that the force and potential at a point within the cube are the same as if it were not there. The resultant magnetization is equivalent (§ 180) to a distribution of magnetism over the two faces normal to the lines of force, with a density equal to kF, positive on the face turned toward the positive direction of the lines of force, and negative on the other face. Let the length of an edge of the cube be denoted by a. Then the quantity of magnetism on each of the two faces is kFa^2. We are to determine the work done by a movement of the cube from one point in a magnetic field to another. To do this we will determine first the work done by the magnetism on one face of the cube during the movement.

Consider a series of equidistant points, designated by 1, 2, 3, . . . n, on a line of force. Represent by V_1, V_2, . . . V_n the potentials of those points, of which V_1 is the greatest and V_n the least value of the potential between the points 1 and n, and suppose the points to be so taken that the differences of potential between any two consecutive ones is indefinitely small. The work done by a quantity of magnetism equal to kFa^2 in moving from the point 1 to the point 2 is $kFa^2(V_1 - V_2)$. If F_1 and F_2 represent the values of F at the points 1 and 2, the average force in the distance between them may be set equal to $\frac{F_1 + F_2}{2}$, and the average quantity of induced magnetization on the face of the cube considered during this movement is $a^2k\frac{F_1 + F_2}{2}$.

The work done is then—

From 1 to 2, $a^2k\frac{F_1 + F_2}{2}(V_1 - V_2)$

$$= a^2k\left(\frac{F_1 V_1}{2} + \frac{F_2 V_1}{2} - \frac{F_1 V_2}{2} - \frac{F_2 V_2}{2}\right).$$

From 2 to 3, . . $= a^2k\left(\frac{F_2 V_2}{2} + \frac{F_3 V_2}{2} - \frac{F_2 V_3}{2} - \frac{F_3 V_3}{2}\right).$

. . . .

From $n - 1$ to n,

$$= a^2k\left(\frac{F_{n-1} V_{n-1}}{2} + \frac{F_n V_{n-1}}{2} - \frac{F_{n-1} V_n}{2} - \frac{F_n V_n}{2}\right).$$

The work done in moving from the point 1 to the point n is the sum of these terms.

To effect the summation we must show that all terms similar to the two terms $\frac{F_3 V_2}{2} - \frac{F_2 V_3}{2}$ will vanish.

The force at the point 2 is the space rate of change of potential at that point, taken with the opposite sign. If d represent the distance between any two consecutive points, we have $F_2 = -\frac{V_3 - V_1}{2d}$ in the limit, as d approaches zero. So also $F_3 = -\frac{V_4 - V_2}{2d}$ in the limit. Using these values, we have for the sum above mentioned

$$\frac{F_3 V_2 - F_2 V_3}{2} = \frac{V_2^2 - V_2 V_4 - V_1 V_3 + V_3^2}{4d}.$$

Now in the limit, as the distances between the points 1, 2, 3, 4, approach zero, we have $V_2^2 = V_1 V_3$, and $V_3^2 = V_2 V_4$; hence the terms considered and all similar ones vanish. The total work done in moving from 1 to n is

$$a^2 k\left(\frac{F_1 V_1}{2} - \frac{F_n V_n}{2}\right).$$

If we bear in mind our assumption that we may neglect the variation of induced magnetism within the cube in any one position, we may express the work done by the movement of the quantity $-ka^2F$ on the opposite face of the cube from a point at the distance a measured along a line of force from the point 1, to a point similarly situated with respect to the point n, by

$$-a^2 k\left(\frac{F_1(V_1 + \Delta V_1)}{2} - \frac{F_n(V_n + \Delta V_n)}{2}\right).$$

In this expression ΔV_1 represents the difference in potential between the point 1 and the point at distance a from it at which the potential is higher than V_1, and ΔV_n represents a similar difference between the point n and the point at distance a from it at which the potential is higher than V_n. The work done by the whole cube in moving from the point 1 to the point n is the sum of the quantities of work done by the quantities of magnetism on its two faces, and is hence equal to

$$-\frac{a^2k}{2}(F_1\Delta V_1 - F_n\Delta V_n).$$

From the relation between force and potential we have, in the limit, as a becomes indefinitely small, $F_1 = \frac{\Delta V_1}{a}$ and $F_n = \frac{\Delta V_n}{a}$, since the distance a is measured in the negative direction. Substituting these values, the expression for the work done by the cube becomes $W = -\frac{a^3k}{2}(F_1^2 - F_n^2)$. The free movement of any system is such as to do work. Hence the cube will move from the point 1 along the line of force toward n if free to do so, in case W is positive.

Two cases may arise depending on the substance of which the cube is composed. We assume the value of k for vacuum as zero. If k for any body be positive, the body is paramagnetic, and W is > 0 when $F_1 < F_n$; the cube moves from a place of weaker to a place of stronger magnetic force. If k be negative, the body is diamagnetic, and W is > 0 when $F_1 > F_n$; the cube moves from a place of stronger to a place of weaker magnetic force.

The subject may be looked at from a different point of view. The coefficient of induced magnetization k is negative in all diamagnetic bodies, but its numerical value is small. It has never been found to be numerically greater than $\frac{1}{4\pi}$ in diamagnetic bodies. In such bodies, therefore, the value of μ, the magnetic permeability, is less than 1, though never negative. When k is 0, μ equals 1, and for paramagnetic bodies μ is greater than 1. The ratio of the force within the substance of which the magnetic permeability is μ to that in vacuum, in which it is supposed to be placed, is $\frac{N}{F} = 1 + 4\pi k = \mu$. If the convention of § 21 be used, by which the strength of a field of force is represented by the number of lines of force passing perpendicularly through unit area, it is evident that when a paramagnetic body in which $\mu > 1$ and $N > F$ is brought into the field, the lines of force are converged into the body. When a diamagnetic body is in the field the lines of force are deflected from it.

As may be easily seen, a paramagnetic body of permeability μ_1 surrounded by a medium also paramagnetic, but of permeability $\mu_2 > \mu_1$, will act relative to the medium as a diamagnetic body. The condition of any body of which the permeability is less than that of the medium in which it is immersed is like that of a weak magnet between the ends of two stronger ones, all three being magnetized in the same direction. The movements of both paramagnetic and diamagnetic bodies may be roughly illustrated by the movements of bodies immersed in water, which rise or sink according as their specific gravities are less or greater than the specific gravity of water.

185. Changes in Magnetic Moment.—When a magnetizable body is placed in a powerful magnetic field, it often receives, temporarily, a more intense magnetization than it can retain when removed. It is said to be *saturated*, or magnetized to *saturation*, when the intensity of its magnetization is the greatest which it can retain when not under the inductive action of other magnets. The coercive force of steel is much greater than that of any other substance; the intensity of magnetization which it can retain is, therefore, relatively very great, and it is hence used for permanent magnets. It is found that the coercive force depends upon the quality and temper of the steel.

Changes of temperature cause corresponding changes in the magnetic moment of a magnet. If the temperature of a magnet be gradually raised, its magnetic moment diminishes by an amount which, for small temperature changes, is nearly proportional to the change of temperature. The magnet recovers its original magnetic moment when cooled again to the initial temperature, provided that the temperature to which it was raised was never very high. If it be raised, however, to a red heat, all traces of its original magnetism permanently disappear. Trowbridge has shown that, if the temperature of a magnet be carried below the temperature at which it was originally magnetized, its magnetic moment also temporarily diminishes.

Any mechanical disturbance, such as jarring or friction, which increases the freedom of motion among the molecules of a magnet, in general brings about a diminution of its magnetic moment. On the other hand, similar mechanical disturbances facilitate the acquisition of magnetism by any magnetizable body placed in a magnetic field.

186. Theories of Magnetism.—It has been shown by mathematical analysis that the facts of magnetic interactions and distribution are consistent with the hypothesis, which we have already made, that the ultimate molecules of iron are themselves magnets, having north and south poles which attract and repel similar poles in accordance with the law of magnetic force. Poisson's theory, upon which most of the earlier mathematical work was based, was that there exist in each molecule indefinite quantities of north and south magnetic fluids, which are separated and moved to opposite ends of the molecule by the action of an external magnetizing force. Weber's view, which is consistent with other facts that Poisson's theory fails to explain, is that each molecule is a magnet, with permanent poles of constant strength, that the molecules of an iron bar are, in general, arranged so as to neutralize one another's magnetic action, but that, under the influence of an external magnetizing action, they are arranged so that their magnetic axes lie more or less in some one direction. The bar is then magnetized. On this hypothesis there should be a limit to the possibible intensity of magnetization, which would be reached when the axes of all the molecules have the same direction. Direct experiments by Joule and J. Müller indicate the existence of such a limit. An experiment of Beetz, in which a thin filament of iron deposited electrolytically in a strong magnetic field becomes a magnet of very great intensity, points in the same direction. The coercive force is, on this hypothesis, the resistance to motion experienced by the molecules. The facts that magnetization is facilitated by a jarring of the steel brought into the magnetic field, that a bar of iron or steel after being removed from the magnetic field retains some of its magnetic properties, that the dimensions of an iron bar are altered by magnetization, the bar becoming longer and diminishing in cross-section, and that a magnetized steel bar loses its magnetism if it be highly heated, are all facts which are best explained by Weber's hypothesis.

CHAPTER II.

ELECTRICITY IN EQUILIBRIUM.

187. Fundamental Facts.—(1) If a piece of glass and a piece of resin be brought in contact, or preferably rubbed together, it is found that, after separation, the two bodies are attracted towards each other. If a second piece of glass and a second piece of resin are treated in like manner, it is found that the two pieces of glass repel each other and the two pieces of resin repel each other, while either piece of glass attracts either piece of resin. These bodies are said to be *electrified* or *charged.*

All bodies may be electrified, and in other ways than by contact. It is sufficient for the present to consider the single example presented. The experiment shows that bodies may be in two distinct and dissimilar states of electrification. The glass treated as has been described is said to be vitreously or *positively electrified*, and the resin resinously or *negatively electrified.* The experiment shows also that bodies similarly electrified repel one another, and bodies dissimilarly electrified attract one another.

(2) If a metallic body, supported on a glass rod, be touched by the rubbed portion of an electrified piece of glass, it will become positively electrified. If it be then joined to another similar body by means of a metallic wire, the second body is at once electrified. If the connection be made by means of a damp linen thread, the second body becomes electrified, but not so rapidly as before. If the connection be made by means of a dry white silk thread, the second body shows no signs of electrification, even after the lapse of a considerable time. Bodies are divided according as they can be classed with the

metals, damp linen, or silk, as *good conductors*, *poor conductors*, and *insulators*. The distinction is one of degree. All conductors offer some opposition to the transfer of electrification, and no body is a perfect insulator.

A conductor separated from all other conductors by insulators is said to be *insulated*. A conductor in conducting contact with the earth is said to be *grounded* or joined to ground.

During the transfer of electrification in the experiment above described the connecting conductor acquires certain properties which will be considered under the head of Electrical Currents.

(3) If a positively electrified body be brought near an insulated conductor, the latter shows signs of electrification. The end nearer the first body is negatively, the farther end positively, electrified. If the first body be removed, all signs of electrification on the conductor disappear. If, before the first body is removed, the conductor be joined to ground, the positive electrification disappears. If now the connection with ground be broken, and the first body removed, the conductor is negatively electrified.

The experiment can be carried out so as to give quantitative results, in a way first given by Faraday. An electrified body, for example a brass ball suspended by a silk thread, is introduced into the interior of an insulated closed metallic vessel. The exterior of the vessel is then found to be electrified in the same way as the ball. This electrification disappears if the ball be removed. If the ball be touched to the interior of the vessel, no change in the amount of the external electrification can be detected. If, after the ball is introduced into the interior, the vessel be joined to ground by a wire, all external electrification disappears. If the ground connection be broken, and the ball removed, the vessel has an electrification dissimilar to that of the ball. If the ball, after the ground connection is broken, be first touched to the interior of the vessel and then removed, neither the ball nor the vessel is any longer electrified.

A body thus electrified without contact with any charged body is said to be electrified by *induction*. The above-mentioned facts show that an insulated conductor, electrified by induction, is electrified both positively and negatively at once, that the electrification of a dissimilar kind to that of the inducing body persists, however the insulation of the conductor be afterwards modified, and that the total positive electrification induced by a positively charged body is equal to that of the inducing body, while the negative electrification can exactly neutralize the positive electrification of the inducing body.

The use of the terms positive and negative is thus justified, since they express the fact that equal electrifications of dissimilar kinds are exactly complementary, so that, if they be superposed on a body, that body is not electrified. These two kinds of electrification may then be spoken of as opposite.

If the glass and resin considered in the first experiment be rubbed together within the vessel, and in general if any apparatus which produces electrification be in operation within the vessel, no signs of any external electrification can be detected. It is thus shown that, whenever one state of electrification is produced, an equal electrification of the opposite kind is also produced at the same time.

Franklin showed that, by the use of a closed conducting vessel of the kind just described, a charged conductor introduced into its interior and brought into conducting contact with its walls is always completely discharged, and the charge is transferred to the exterior of the vessel. This procedure furnishes a method of adding together the charges on any number of conductors, whether they be charged positively or negatively. It is thus theoretically possible to increase the charge of such a conductor indefinitely.

(4) If any instrument for detecting forces due to electrifications be introduced into the interior of a closed conductor charged in any manner, it is found that no signs of force due to the charge can be detected. The experiment was accurately executed by Cavendish, and afterwards tried on a large scale by Faraday. It proves that within a closed electrified conductor there is no electrical force due to the charge on the conductor, or that the potential due to the electrical forces is uniform within the conductor.

188. Law of Electrical Force.—If two charged bodies be considered, of dimensions so small that they may be neglected in comparison with the distance between the bodies, the stress between the two bodies due to electrical force is proportional directly to the product of the charges which they contain, and inversely to the square of the distance between them.

If Q and $Q_{,}$ represent two similar charges, r the distance between them, and k a factor depending on the units in which the charges are measured, the formula expressing the repulsion between them is

$$k\frac{QQ_{,}}{r^2}.$$

Coulomb used the torsion balance (§ 82) to demonstrate this law. At one end of a glass rod suspended from the torsion wire and turning in the horizontal plane is placed a gilded pith ball, and through the lid of the case containing the apparatus can be introduced a similar insulated ball so arranged that its centre is at the same distance from the axis of rotation of the suspended system, and in the same horizontal plane, as the centre of the first ball. This second ball may be called the *carrier*.

To prove the law as respects quantities, the suspended ball is brought into equilibrium at the point afterwards to be occupied by the carrier ball. The carrier ball is then charged and introduced into the case. When it comes in contact with the suspended ball, it shares its charge with it and a repulsion ensues. The torsion head must then be rotated until the suspended ball is brought to some fixed point, at a distance from the carrier which is less than that which would separate the two balls in the second part of the experiment if no torsion were brought upon the wire. The repulsion is then measured in terms of the torsion of the wire. The charge on the carrier is then halved, by touching it with a third similar insulated ball, and, the charge on the suspended ball remaining the same, the repulsion between the two balls at the same distance is again observed. If the case be so large that no disturbing effect of the walls enters, and if the balls be small and so far apart that their inductive action on one another may be neglected, the repulsion in the second case is found to be one half that in the first case. In general the problem is a far more difficult one, for the distribution on the two spheres is not uniform. That portion of the distribution dependent on the induction of the balls can be calculated, but the irregularities of distribution due to the action of the walls of the case and other disturbing elements can only be allowed for approximately.

The law as respects distance is proved in a somewhat similar way. The repulsions at two different distances are measured in terms of the torsion of the wire, the charges on the two balls remaining the same. The same corrections must be introduced as in the former case.

189. Distribution.—The law of electrical force has been stated in terms of the charges of two bodies. We may, however, consider electricity as a quantity which has an existence independent of matter and which is distributed in space. The fact cited in § 187 (4) shows that this distribution must be looked on as being on the surfaces of conductors and not on their interiors. If we define *surface density* of electrification at any point on the surface of a charged conductor as the limit of the ratio of the quantity of electricity on an element in the surface at that point to the area of the element as that area approaches zero, we may measure quantities of electricity in terms of surface density. The surface density of electricity is usually designated by σ.

If the law of electrical force hold true not only for charges on bodies but also for quantities of electricity on the surface elements of a conductor, it is evident, from the fact that within an electrified conductor there is no electrical force, that its surface density of electrification must be proportional at every point on its surface to the thickness at that point of a shell of matter which is so distributed on that surface that there is no force at any point enclosed by the surface. The distribution on a charged sphere may, from symmetry, be assumed uniform. The fact that there is no electrical force within a charged sphere is then, from § 29 (1), consistent with the law of electrical force which has been given; and since the means of detecting electrical force, if there were any, within a charged conductor are very delicate, this fact affords a strong corroborative proof of the law.

The determination of the distribution of electricity on irregularly shaped conductors is in general beyond our power. If we consider, however, a conductor in the form of an elongated egg, it can be readily seen that, in order that there may be no electrical force within it, the surface density at the pointed end must be greater than that anywhere else on its surface. In general, the surface density at points on a conducting surface depends upon the curvature of the surface, being greater where the curvature is greater. Thus, if the conductor be a long rod terminating in a point, the surface density at the pointed end is much greater than that anywhere else on the rod.

190. Unit Charge.—The law of electrical force enables us to define a *unit charge*, based upon the fundamental mechanical units.

Let there be two equal and similar positive charges concentrated at points unit distance apart in air, such that the repulsion between them equals the unit of force. Then each of the charges is a *unit charge*, or a *unit quantity of electricity*. With this definition of unit charge, it may be said that the force between two charges is not merely proportional to, but equals, the product of the charges divided by the square of the distance between them. The factor k in the expression for the force between two charges becomes unity, and the dimensions of $\frac{QQ_{\prime}}{r^2}$ are those of a force. If the charges be equal, we have

$$\left[\frac{Q^2}{r^2}\right] = MLT^{-2}.$$

Hence $[Q] = M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-1}$ are the dimensions of the charge. This equation gives the charge in absolute mechanical units, and by means of it all other electrical quantities may be expressed in absolute units. It is at the basis of the *electrostatic system* of electrical measurements.

The practical unit of charge or quantity is called the *coulomb*. It is the quantity of electricity transferred during one second by a current of one ampere (§ 218).

191. Electrical Potential.—The electrical forces have a potential similar to that discussed in § 28. The unit quantity of positive electricity is taken as the test unit. Since [§ 187 (4)] the potential at every point of a charged conductor is the same, the surface of the conductor is an equipotential surface. The potential of this surface is often called the *potential of the conductor*. A conductor joined to ground is at the potential of the earth. It will be shown (§ 195) that the potential of the earth is not appreciably modified when a charged conductor is joined to ground. All conductors, moreover, however they may afterwards be charged, are when uncharged at the potential of the earth. For these reasons it is usual to take the potential of the earth as the fixed potential or zero from which to reckon the potentials of electrified bodies. The potential of a freely electrified conductor and of the region about it is thus positive when the charge of the conductor is positive, and negative when it is negative. A conductor joined to ground is at zero potential.

The difference of potential between two points is equal to the work done in carrying a unit quantity of electricity from one point to another. We then have the equation $Q(V_{\prime} - V)$ = work. Hence follows the dimensional equation $[V_{\prime} - V] = \frac{ML^2T^{-2}}{M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-1}} = M^{\frac{1}{2}}L^{\frac{1}{2}}T^{-1}$, the dimensions of difference of potential in electrostatic units.

If any distribution of a charge exist on a conductor, which is such that the potential at all points in the conductor is not the same, it is unstable, and a rearrangement goes on until the potential becomes everywhere the same. The process of rearrangement is said to consist in a flow of electricity from points of higher to points of lower potential.

On this property of electricity depends the fact that a closed conducting surface completely screens bodies within it from the action of external electrical forces. For, whatever changes in potential occur in the region outside the closed conductor, a redistribution will take place in it such as to make the potential of every point within it the same. Electrical force depends on the space rate of change of potential, and not on its absolute value. Hence the changes without the closed conductor will have no effect on bodies within it. Further, any electrical operations whatever within the closed conductor will not change the potential of points outside it. For, whatever operations go on, equal amounts of positive and negative electricity always exist within the conductor, and hence the potential of the conductor remains unaltered. Hence electrical experiments performed within a closed room yield results which are as valid as if the experiments were performed in free space.

The advantage gained by the use of the idea of potential in discussions of electrical phenomena may be illustrated by a statement of the process of charging a conductor by induction described in § 187 (3). To fix our ideas, let us suppose that the field of force is due to a positively electrified sphere, and that the body to be charged is a long cylinder. When this cylinder, previously in contact with the earth and therefore at zero potential, is brought end on to a point near the sphere, it is in a region of positive potential, and is itself at a positive potential. If we consider the original potentials at the points in the region now occupied by the cylinder, it is easily seen that the potential of points nearer the sphere was higher than that of those more remote. When the cylinder is brought into the field, therefore, the portion nearer the sphere is temporarily raised to a higher potential than the portion more remote. The difference of potential between these portions is annulled by a flow of electricity from the points of higher potential to those of lower potential at a rate depending on the conductivity of the cylinder. The end of the cylinder nearer the sphere is negatively charged, the end more remote is positively charged, and the two charged portions are separated by a line on the surface, called the neutral line, on which there is no charge.

If the cylinder be now joined to ground, a flow of electricity takes place through the ground connection, and it is brought to zero potential. The potential of the cylinder is therefore everywhere lower than the original potentials of the points in the region which it occupies. This necessitates a negative charge distributed over the whole cylinder. In other words, the earth and the cylinder may be considered as forming one conductor charged by induction, in which the neutral line is not within the cylinder.

If the ground connection be broken the electrical relations are not disturbed. If the cylinder be now removed to a region of lower potential against the attraction of the sphere, work will be done against electrical forces, which reappears as electrical energy. The potential of the cylinder is lowered, and, if it be again connected with the earth, work will be done by a flow of electricity to it.

The fact that there is no electrical force within a closed electrified conductor of any shape permits some extensions of the theorems of § 29.

Some small portion of the surface of any electrified conductor may be considered a plane relatively to a point situated just outside it. Represent the surface density of electricity on that plane by σ. It was proved (§ 29) that the force due to such a plane is $2\pi\sigma$, if we substitute σ for the corresponding factor *d*. Now, just inside the conductor the force is zero. This results from the equilibrium of the force due to the plane portion and that due to the rest of the conductor. The force due to the rest of the conductor is therefore $2\pi\sigma$. At a point just outside the conductor these two forces act in the same direction. Hence the total force due to the conductor at a point just outside it is the sum of the two forces, or $4\pi\sigma$.

From the preceding proposition follows at once a deduction as to the pressure outwards on the surface of an electrified conductor due to the repulsion of the various parts of the charge for one another. Select any small portion of the surface of the electrified conductor of area a. The force on unit quantity acting outward from the conductor at a point in that area due to the charge of the rest of the conductor is $2\pi\sigma$. This force acts on every unit of charge on the area. The force on the area acting outwards is then $2\pi a\sigma^2$, or the pressure at a point in the area referred to unit of area is $2\pi\sigma^2$. This quantity is often called the *electric pressure*.

192. Capacity.—The electrical *capacity* of a conductor is defined to be the charge which the conductor must receive to raise it from zero to unit potential, while all other conductors in the field are kept at zero potential. This charge varies for any one conductor in a way which cannot be always definitely determined, depending upon the medium in which the conductor is immersed and the position of other conductors in the field. When the charged conductor is in very close proximity to another conductor which is kept at zero potential, the amount of charge needed to raise it to unit potential is very great as compared with that required when the other conductor is more remote. Such an arrangement is called a *condenser*. If the charge on a conductor be increased, the increase in potential is directly as that of the charge. Hence the capacity C is given by dividing any given charge on a conductor by the potential of that conductor, or

$$C = \frac{Q}{V}. \tag{78}$$

The practical unit of capacity is the *farad*, which is the capacity of a conductor, the charge on which is one coulomb (§ 190) when its potential is one volt (§ 228). This unit is too great for convenient use. Instead of it a *microfarad*, or the one-millionth part of a farad, is usually employed.

This equation gives the dimensions of capacity. Measured in electrostatic units, they are

$$[C] = \left[\frac{Q}{V}\right] = \frac{M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-1}}{M^{\frac{1}{2}}L^{\frac{1}{2}}T^{-1}} = L.$$

Capacity, therefore, is of the dimensions of a length.

In the theory of Faraday, which has been adopted and developed by Maxwell, electrification is made to consist in an arrangement or displacement of the insulating medium, called by him the *dielectric*, surrounding the electrified conductor.

This displacement, beginning at the surface of the electrified conductor, continues throughout the dielectric until it terminates at the surfaces of other conductors. The electrification of the charged conductor is the manifestation of this displacement at one face of the dielectric, that of the surrounding conductors the manifestation of the displacement on the other face. The one charge cannot exist without an equal and opposite charge on surrounding conductors, as was experimentally proved by Faraday's experiment already described in § 187 (3). It is therefore necessary, in considering the capacity of any conductor, to take account of the medium in which it is immersed, and of the arrangement of surrounding conductors.

193. Specific Inductive Capacity.—The fact that the capacity of a condenser of given dimensions depends upon the medium used as the dielectric was first discovered by Cavendish, and afterwards rediscovered by Faraday. The property of the medium upon which this fact depends is called its *specific inductive capacity*. The specific inductive capacity of vacuum is taken as the standard. If Q represent the charge required to raise a condenser in which the dielectric is vacuum to a potential V, then if another dielectric be substituted for vacuum, it is found that a different charge $Q_{\prime}$ is required to raise the potential to V. The ratio $\frac{Q_{\prime}}{Q} = K$ is the specific inductive capacity. Since $C_{\prime} = \frac{Q_{\prime}}{V}$ and $C = \frac{Q}{V}$ are the capacities of the condenser with the two dielectrics, it follows that

$$C_{\prime} = CK, \tag{79}$$

where C is the capacity with vacuum as the dielectric. The specific inductive capacity K is always greater than unity. Some dielectrics, such as glass and hard rubber, have a high specific inductive capacity, and at the same time are capable of resisting the strain put upon them by the electric displacement to a much greater extent than such dielectrics as air. They are therefore used as dielectrics in the construction of condensers.

194. Condensers.—The simplest condenser, one which admits of the direct calculation of its capacity, and from which the capacities of many other condensers may be approximately calculated or inferred, consists of a conducting sphere surrounded by another hollow concentric conducting sphere which is kept always at zero potential by a ground connection. For convenience we assume the specific inductive capacity of the dielectric separating the spheres to be unity.

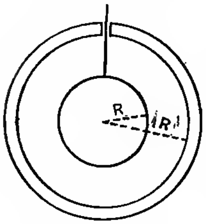

FIG. 58.

Let the radius of the small sphere (Fig. 58) be denoted by R, that of the inner spherical surface of the larger one by $R_{\prime}$; let a charge Q be given to the inner sphere by means of a conducting wire passing through an opening in the outer sphere, which may be so small as to be negligible. This charge Q will induce on the outer sphere an equal and opposite charge, $-Q$. Since the distribution on the surface of the spheres may be assumed uniform, the potential at the centre of the two spheres, due to the charge on the inner one, is $\frac{Q}{R}$, and the potential due to the charge of the outer sphere is $-\frac{Q}{R_{\prime}}$. Hence the actual potential V at the centre, due to both charges, is

$$\frac{Q}{R} - \frac{Q}{R_{\prime}} = Q\left(\frac{R_{\prime} - R}{RR_{\prime}}\right).$$

Hence the capacity is

$$C=\frac{Q}{V}=\frac{RR}{R_{,}-R}. \tag{80}$$

In order to find the effect of a variation of the value of R, divide numerator and denominator by $R_{,}$ and write

$$C=\frac{R}{1-\frac{R}{R_{,}}}.$$

Now, if $R_{,}$ be greater than R by an infinitesimal, the fraction $\frac{R}{R_{,}}$ is less than unity by an infinitesimal, and the capacity of the accumulator is infinitely great. It becomes infinitely small if R be diminished without limit. The presence of any finite charge at a point would require an infinite potential at that point, which is of course impossible. The existence of finite charges concentrated at points, which we have assumed sometimes in order to more conveniently state certain laws, is therefore purely imaginary. If electricity is distributed in space, it is distributed like a fluid, a finite quantity of which never exists at a point.

If $R_{,}$ increase without limit, C becomes more and more nearly equal to R. Suppose the inner sphere to be surrounded not by the outer sphere but by conductors disposed at unequal distances, the nearest of which is still at a distance $R_{,}$ so great that $\frac{R}{R_{,}}$ may be neglected in comparison with unity. Then if the nearest conductor were a portion of a sphere of radius $R_{,}$ concentric with the inner sphere, the capacity of the inner sphere would be approximately R. And this capacity is evidently not less than that which would be due to any arrangement of conductors at distances more remote than $R_{\prime}$. Therefore the capacity of a sphere removed from other conductors by distances very great in comparison with the radius of the sphere is equal to its radius R. This value R is often called the capacity of a *freely electrified* sphere. Strictly speaking, a freely electrified conductor cannot exist; the term is, however, a convenient one to represent a conductor remote from all other conductors.

A common form of condenser consists of two flat conducting disks of equal area, placed parallel and opposite one another. The capacity of such a condenser may be calculated from the capacity of the spherical condenser already discussed. Let d represent the distance $R_{\prime} - R$ between the two spherical surfaces. Let A and $A_{\prime}$ represent the area of the surfaces of the two spheres of radius R and $R_{\prime}$. Then we have

$$R^2 = \frac{A}{4\pi} \quad \text{and} \quad R_{\prime}^2 = \frac{A_{\prime}}{4\pi}.$$

The capacity of the spherical condenser may then be written

$$\frac{\sqrt{AA_{\prime}}}{4\pi d}.$$

If $R_{\prime}$ and R increase indefinitely, in such a manner that $R_{\prime} - R$ always equals d, in the limit the surfaces become plane and A becomes equal to $A_{\prime}$. The capacity therefore equals $\frac{A}{4\pi d}$. Since the charge is uniformly distributed, the capacity of any portion of the surface cut out of the sphere is proportional to the area S of that surface, or

$$C = \frac{S}{4\pi d}. \tag{81}$$

This value is obtained on the assumption that the distribution over the whole disk is uniform, and the irregular distribution at the edges of the disk is neglected. It is therefore only an approximation to the true capacity of such a condenser.

The so-called *Leyden jar* is the most usual form of condenser in practical use. It is a glass jar coated with tinfoil within and without, up to a short distance from the opening. Through the stopper of the jar is passed a metallic rod furnished with a knob on the outside and in conducting contact with the inner coating of the jar. To charge the jar, the outer coating is put in conducting contact with the ground, and the knob brought in contact with some source of electrification. It is discharged when the two coatings are brought in conducting contact. When the wall of the jar is very thin in comparison with the diameter and with the height of the tinfoil coating, the capacity of the jar may be inferred from the preceding propositions. It is approximately proportional directly to the coated surface, to the specific inductive capacity of the glass, and inversely to the thickness of the wall.

195. Systems of Conductors.—If the capacities and potentials of two or more conductors be known, the potential of the system formed by joining them together by conductors is easily found. It is assumed that the connecting conductors are fine wires, the capacities of which may be neglected. Then the charges of the respective bodies may be represented by C_1V_1, C_2V_2, . . . C_nV_n, and the capacity of the system by the sum $C_1 + C_2 + \ldots C_n$. Hence V_1, the potential after connections have been made, is

$$V = \frac{C_1V_1 + C_2V_2 + \ldots C_nV_n}{C_1 + C_2 + \ldots C_n}$$

In the case of two freely electrified spheres joined up together by a fine wire, we have $C_1 = R_1$, and $C_2 = R_2$, where R_1 and R_2 represent the radii of the spheres. Hence we have

$$V = \frac{R_1V_1 + R_2V_2}{R_1 + R_2}.$$

When R_1 is very great compared with R_2, we obtain

$$V = V_1 + \frac{R_2}{R_1}V_2.$$

Unless V_2 is so great that the term $\frac{R_2}{R_1}V_2$ becomes appreciable, the potential of the system is appreciably equal to the original potential of the larger sphere. Manifestly the same result follows if R_2 represent the capacity of any conductor relatively small compared with the capacity of the large sphere. This proposition justifies the adoption of the potential of the earth as the standard or zero potential.

196. Energy of Charge.—In order to find the work done in charging a conducting body to a given potential, we will consider all surrounding bodies as being kept by ground connections at zero potential. Then if an infinitesimal charge be given to the body, previously uncharged and at zero potential, the work done is that which would be done if the charge were brought from infinity to a point of potential 0; that is, the work = 0. The charge q raises the potential of the body so that it becomes $v_1 = \frac{q}{C}$. If then another infinitesimal charge q be given to the body, the work done is equal to qv_1 or $\frac{q^2}{C}$, and the potential is raised to $v_2 = \frac{2q}{C}$. So also the work done when the $(n+i)$th charge is given to the body is qv_n, and the potential becomes $\frac{(n+i)q}{C}$. The total work done is then

$$W = q(v_1 + v_2 + \ldots v_n) = \frac{q^2}{C}(1 + 2 + \ldots n)$$

$$= \frac{(n+1)n}{2} \cdot \frac{q^2}{C} = \tfrac{1}{2}QV, \tag{82}$$

where $nq = Q$ and $V = v_n$. When the charges q are infinitesimal, Q is equal to the sum of all the charges given to the body. Hence the work done in raising a body from zero potential to potential V is equal to one half the charge multiplied by the potential of the body.

197. Strain in the Dielectric.—An instructive experiment illustrating Faraday's theory that the electrification of a conductor is due to an arrangement in the dielectric surrounding it, may be performed with a jar so constructed that both coatings can be removed from it. If the jar be charged, the coatings removed by insulating handles without discharging the jar, and examined, they will be found to be almost without charge. If they be replaced, the jar will be found to be charged as before. The jar will also be found to be charged if new coatings similar to those removed be put in their place. This result shows that the true seat of the charge is in the dielectric. The experiment is due to Franklin.

That the arrangement in the dielectric is of the nature of a strain is rendered probable by the fact, first noticed by Volta, that the volume occupied by a Leyden jar increases slightly when the jar is charged. Similar changes of volume were observed by Quincke in fluid dielectrics as well as in different solids.

Another proof of the strained condition of dielectrics is found in their optical relations. It was discovered by Kerr that dielectrics previously homogeneous become doubly refracting when subjected to a powerful electrical stress. Maxwell has shown, from the assumptions of his electromagnetic theory of light, that the index of refraction of a transparent dielectric should be proportional to the square root of its specific inductive capacity. Numerous experiments, among which those of Boltzmann on gases are the most striking, show that this predicted relation is very close to the truth.

It has further been shown that the specific inductive capacity of sulphur has different values along its three crystallographic axes. This is probably true also for other crystals.

Some crystals, while being warmed, exhibit on their faces positive and negative electrifications, which are reversed as the crystals are cooling. This fact, while as yet unexplained, is probably due to temporary modifications of molecular arrangement by heat.

If a jar be discharged and allowed to stand for a while, a second discharge can be obtained from it. By similar treatment several such discharges can be obtained in succession. The charge which the jar possesses after the first discharge is called the *residual charge*. It does not attain its maximum immediately, but gradually, after the first discharge. The attainment of the maximum is hastened by tapping on the wall of the jar. This phenomenon was ascribed by Faraday to an absorption of electricity by the dielectric, but this explanation is at variance with Faraday's own theory of electrification. Maxwell explains it by assuming that want of homogeneity in the dielectric admits of the production of induced electrifications at the surfaces of separation between the non-homogeneous portions. When the jar is discharged the induced electrifications within the dielectric tend to reunite, but, owing to the want of conductivity in the dielectric, the reunion is gradual. After a sufficient time has elapsed, the alteration of the electrical state of the dielectric has proceeded so far as to sensibly modify the field outside the dielectric. The residual charge then appears in the jar.

198. Electroscopes and Electrometers.—An *electroscope* is an instrument to detect the existence of a difference of electrical potential. It may also give indications of the amount of difference. It consists of an arrangement of some light body or bodies, such as a pith ball suspended by a silk thread, or a pair of parallel strips of gold-foil, which may be brought near or in contact with the body to be tested. The movements of the light bodies indicate the existence, nature, and to some extent the amount of the potential difference between the body tested and surrounding bodies.

An *electrometer* is an apparatus which gives precise measurements of differences of potential. The most important form is the *absolute* or *attracted disk electrometer*, originally devised by Harris, and improved by Thomson. The essential portions of the instrument (Fig. 59) are a large flat disk B which can be put in conducting contact with one of the two bodies between which the difference of potential is desired; a similar disk C, in the centre of which is cut a circular opening, placed parallel to and a little distance above the former one; a smaller disk A with a diameter a little less than that of the opening, which can be placed accurately in the opening and brought plane with the larger disk; and an arrangement, either a balance arm or a spring of known strength, from which the small disk is suspended, and by means of which the force acting on the disk when it is plane with the surface of the larger disk can be measured. The three disks can be conveniently styled the *attracting disk*, the *guard ring*, and the *attracted disk*. The position of the attracted disk when it is in the plane of the guard ring is often called the *sighted position*. The guard ring is employed in order that the distribution on the attracted disk may be uniform.

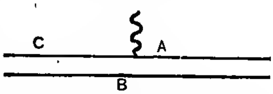

FIG. 59.

To determine the difference of potential between the attracted and attracting disks, we consider them first as forming a flat condenser. If we represent by Q the quantity of electricity on the attracted disk, by V and V_1 the potentials of the attracted and attracting disks respectively, by d the distance between them, and by S the area of the attracted disk, then, as has been shown in § 194, the capacity of such a condenser is

$$\frac{Q}{V_1 - V} = \frac{S}{4\pi d}.$$

Now from the nature of the condenser, and in consequence of the regular distribution due to the presence of the guard ring, we have $\frac{Q}{S} = \sigma$, the surface density on either plate, whence $\sigma = \frac{V_1 - V}{4\pi d}$. The surface density cannot be measured, and must be eliminated by means of an equation obtained by observation of the force with which the two disks are attracted. The plates are never far apart, and the force on a unit charge due to the charge on the lower one may be always taken in the space between the plates as equal to $2\pi\sigma$ (§ 191). Every unit on the attracted disk is attracted with this force, and the total attraction, which is measured by means of the balance or spring, is $F = 2\pi\sigma^2 S$. Substituting this value of σ in the former equation, we get

$$V_1 - V = d\sqrt{\frac{8\pi F}{S}}, \qquad (83)$$

which gives the difference of potential between the two plates in terms which are all measurable in absolute units. In Thomson's form of the electrometer the attracted disk is kept at a high constant potential V; the attracting disk is brought to the potential V_1 of one of the two bodies of which the difference of potential is desired, and the position of the attracting disk when the attracted disk is in its sighted position is noted. The attracting disk is then brought to the potential V_2 of the other body, and by a micrometer screw the distance is measured through which the attracting disk is moved in order to bring the attracted disk again into its sighted position. This measurement can be made with much greater precision than the measurement of the distance between the two plates. The formula is easily deduced from the one already given. In the first observation we have

$$V_1 - V = d_1\sqrt{\frac{8\pi F}{S}};$$

in the second,

$$V_2 - V = d_2\sqrt{\frac{8\pi F}{S}};$$

whence

$$V_1 - V_2 = (d_1 - d_2)\sqrt{\frac{8\pi F}{S}}, \qquad (84)$$

and $d_1 - d_2$ is the distance measured.

Thomson's *quadrant electrometer* is an instrument which is not used for absolute measurement, but being extremely sensitive to minute differences of potential, it enables us to compare them with each other and with some known standard. The construction of the apparatus can best be understood from

Fig. 60. Of the four metallic quadrants which are mounted on insulating supports, the two marked P and the two marked N are respectively in conducting contact by means of wires. The body C, technically called the *needle*, is a thin sheet of metal, suspended symmetrically just above the quadrants by two parallel silk fibres, forming what is known as a bifilar suspension. When there is no charge in the apparatus, the axes of symmetry of the needle lie above the spaces which separate the quadrants.

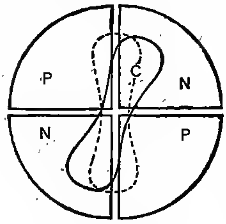

FIG. 60.

To use the apparatus, the needle is maintained at a high, constant potential, and the two points, the difference of potential between which is desired, are joined to the pairs of quadrants P and N. The needle is deflected from its normal position, and the amount of deflection is an indication of the difference of potential between the two pairs of quadrants.

199. **Electrical Machines.**—*Electrical machines* may be divided into two classes: those which depend for their operation upon friction, and those which depend upon induction.

The *frictional machine*, in one of its forms, consists of a circular glass *plate*, mounted so that it can be turned about an axis, and a *rubber* of leather, coated with a metal amalgam, pressed against it. The rubber is mounted on an insulating support, but, during the operation of the machine, it is usually joined to ground. Diametrically opposite is placed a row of metal points, fixed in a metallic support, constituting what is technically called the *comb*. The comb is usually joined to an accessory part of the machine presenting an extended metallic surface, called the *prime conductor*. The prime conductor is carried on an insulating support.

When the plate is turned, an electrical separation is produced by the friction of the rubber, and the rubbed portion of the plate is charged positively. When the charged portion of the plate passes before the comb, an electrical separation occurs in the prime conductor due to the inductive action of the plate, a negative charge passes from the comb to neutralize the positive charge of the plate, and the prime conductor is charged positively. Since accessions are received to the charge of the prime conductor as each portion of the plate passes the comb, it is evident that the potential of the prime conductor will continuously rise, until it is the same as that of the plate, or until a discharge takes place.

The fundamental operations of all *induction machines* are presented by the action of the *electrophorus*, an instrument invented by Volta in 1771. It consists of a plate of sulphur or rubber, which rests on a metallic plate, and a metallic disk mounted on an insulating handle. The sulphur is electrified negatively by friction, and the disk, placed upon it and joined to ground, is charged positively by induction. When the ground connection is broken and the disk lifted from the sulphur, its positive charge becomes available. The process is precisely similar to that described in § 191. It may evidently be repeated indefinitely, and the electrophorus may be used as a permanent source of electricity.

It is evident that a charged metallic plate may be substituted for the sulphur in the construction of an electrophorus, provided that the disk be not brought in contact with it, but only near it. A plan by which this is realized, and at the same time an imperceptible charge on one plate is made to develop an indefinite quantity of electricity of high potential, is shown in Fig. 61. A_1 and A_2 are conducting plates, called *inductors*. In front of them

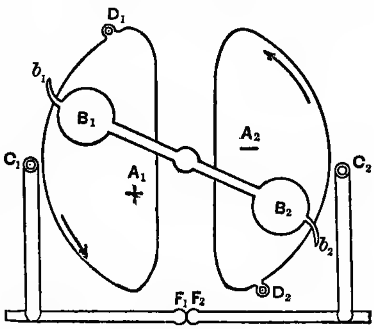

FIG. 61.

two disks B_1 and B_2, called *carriers*, are mounted on an arm so as to turn about the axis E. Projecting springs b_1 and b_2 attached to these disks are so fixed as to touch successively the pins D_1 and D_2, connected with the plates A_1 and A_2, and the pins C_1 and C_2, insulated from the plates, but joined to the prime conductors F_1 and F_2.

Suppose the prime conductors to be in contact and the carriers so placed that B_1 is between D_1 and C_1, and suppose the plate A_1 to be at a slightly higher potential than the rest of the machine. The carrier B_1 is then charged by induction. When the carriers are turned in the direction of the arrows, and the carrier B_1 makes contact with the pin C_1, it loses a part of its positive charge and the prime conductors become positively charged. At the same time the carrier B_2 becomes positively charged. As the carrier B_2 passes over the upper part of the plate A_2, the lower part of the plate A_2 is charged positively by induction. This positive charge is neutralized by the negative charge of the carrier B_1, when contact is made at D_2. The plate A_2 is then negatively charged. The carrier B_2 at its contact at D_1 shares its positive charge with the plate A_1. The carriers then return to the positions from which they started, and the difference of potential between the plates A_1 and A_2 is greater than it was at first. When, after sufficient repetition of this process, the difference of potential has become sufficiently great, the prime conductors may be separated, and the transfer of electricity between the points F_1 and F_2 then takes place through the air. Obviously the number of carriers may be increased, with a corresponding increase in the rapidity of action of the machine. This improvement is usually effected by attaching disks of tin-foil at equal distances from each other on one face of a glass wheel, so that, as the wheel revolves, they pass the contact points in succession.

Another induction machine, invented by Holtz, differs in plan from the one just described in that the metallic carriers, are replaced by a revolving glass plate, and the two metallic inductor plates, by a fixed glass plate. In the fixed plate are cut two openings, diametrically opposite. Near these openings, and placed symmetrically with respect to them, are fixed upon the back of the plate two paper sectors or *armatures*, terminating in points which project into the openings. In front of the revolving plate and opposite the ends of the armatures nearest the openings are the combs of two prime conductors. Opposite the other ends of the armatures, and also in front of the revolving wheel, are two other combs joined together by a cross-bar.

In order to set this machine in operation, one of the paper armatures must be charged from some outside source. The surface of the revolving plate performs the functions of the carriers in the induction machine already explained. The armatures take the place of the inductors, and the points in which they terminate serve the same purpose as the contact points in connection with the inductors. The explanation of the action of this machine is, in general, similar to that already given. The effect of the combs joined by the cross-bar is equivalent to joining to ground that portion of the outside face of the revolving plate which is passing under them.

CHAPTER III.

THE ELECTRICAL CURRENT.

200. Fundamental Effects of the Electrical Current.—In 1791 Galvani of Bologna published an account of some experiments made two years before, which opened a new department of electrical science. He showed that, if the lumbar nerves of a freshly skinned frog be touched by a strip of metal and the muscles of the hind leg by a strip of another metal, the leg is violently agitated when the two pieces of metal are brought in contact. Similar phenomena had been previously observed, when sparks were passing from the conductor of an electrical machine in the vicinity of the frog preparation.

He ascribed the facts observed to a hypothetical animal electricity or vital principle, and discussed them from the physiological standpoint; and thus, although he and his immediate associates pursued his theory with great acuteness, they did not effect any marked advance along the true direction. Volta at Pavia followed up Galvani's discovery in a most masterly way. He showed that, if two different metals, or, in general, two heterogeneous substances, be brought in contact, there immediately arises a difference of electrical potential between them. He divided all bodies into two classes. Those of the *first class*, comprising all simple bodies and many others, are so related to one another that, if a closed circuit be formed of them or any of them, the sum of all the differences of potential taken around the circuit in one direction is equal to zero. If a body of the *second class* be substituted for one of the first class, this statement is no longer true. There exists then in the circuit a preponderating difference of potential in one direction. Volta described in 1800 his famous voltaic battery. He placed in a vessel, containing a solution of salt in water, plates of copper and zinc separated from one another. When wires joined to the copper and zinc were tested, they were found to be at different potentials, and they could be used to produce the effects observed by Galvani. The effects were heightened, and especially the difference of potential between the two terminal wires was increased, when several such cups were used, the copper of one being joined to the zinc of the next so as to form a series. This arrangement was called by Volta the galvanic battery, but is now generally known as the voltaic battery.

Volta observed that, if the terminals of his battery were joined, the connecting wire became heated.

Soon after Volta sent an account of the invention of his battery to the Royal Society, Nicholson and Carlisle observed that, when the terminals of the battery were joined by a column of acidulated water, the water was decomposed into its constituents, hydrogen and oxygen.

In 1820 Oersted made the discovery of the relation between electricity and magnetism. He showed that a magnet brought near a wire joining the terminals of a battery is deflected, and tends to stand at right angles to the wire. His discovery was at once followed up by Ampère, who showed that, if the wire joining the terminals be so bent on itself as to form an almost closed circuit, and if the rest of the circuit be so disposed as to have no appreciable influence, the magnetic potential at any point outside the wire will be the same as that of a uniform magnetic shell.

In 1834 Peltier showed that, if the terminals of the battery be joined by wires of two different metals, there is a production or an absorption of heat at the point of contact of the wires, depending upon which of the wires is joined to the terminal the potential of which is positive with respect to the other. This fact is referred to as the *Peltier effect.*

201. Electromotive Force.—In 1833 Faraday showed conclusively that if a Leyden jar be discharged through a circuit, it will produce the same thermal, chemical, and magnetic effects as those just described as produced by the voltaic battery.

We know that, in the discharge of a jar, a charge of electricity is transferred from a point at a higher potential to one at a lower. It is reasonable, therefore, to suppose the phenomena under consideration to be also due, in some way, to the transfer of electricity from a higher to a lower potential. Since these phenomena continue without interruption while the circuit is joined up, it is necessary to assume that the voltaic battery maintains a permanent difference of potential. This power of maintaining a difference of potential is ascribed to an *electromotive force* existing in the circuit.

In an actual circuit containing a voltaic battery, if two points on the circuit outside the battery be tested by an electrometer, a difference of potential between them will be found. If the circuit be broken between the two points considered, the difference of potential between them becomes greater. This maximum difference of potential is the sum of finite differences of potential supposed to be due to molecular interactions at the surfaces of contact of different substances in the circuit, and is the measure of the electromotive force. An electromotive force may exist in a circuit in which there are no differences of potential. These cases will be considered later. It is sufficient for the present to consider two points between which a difference of potential is maintained, and which are connected by conductors of any kind whatever.

The dimensions of electromotive force in the electrostatic system are those of difference of potential, or $[E] = M^{\frac{1}{2}} L^{\frac{1}{2}} T^{-1}$.

202. Electrostatic Unit of Current.—Let us denote the potentials at the two points 1 and 2 in the circuit by V_1 and V_2, and let V_1 be greater than V_2; then if, in the time t, a quantity of electricity equal to Q passes through a conductor joining those points from potential V_1 to potential V_2, the amount of work done by it is $Q(V_1 - V_2)$.

If the conductor be a single homogeneous metal or some analogous substance, and no motion of the conductor or of any external magnetic body take place, the whole work done is expended in heating the conductor. If we suppose the transfer to be such that equal quantities of heat are developed in equal times, we may represent the heat produced in the time t by Ht, if H represent the heat developed in one unit of time. If all the quantities considered are expressed in terms of the same fundamental units, we have

$$Q(V_1 - V_2) = Ht, \quad \text{or} \quad H = \frac{Q}{t}(V_1 - V_2).$$

The transfer of electricity in the circuit is called the *electrical current*, and the rate of transfer $\frac{Q}{t} = I$ is called the *current strength*, or often simply the *current*. The current, as here defined, is independent of the nature of the conductor, and is the same for all parts of the circuit. This fact was experimentally proved by Faraday. Employing this quantity I, we have the fundamental equation

$$H = I(V_1 - V_2). \tag{85}$$

If heat and difference of potential be measured in absolute units, this equation enables us to determine the absolute *unit of current*. The system of units here used is the electrostatic system. The dimensions of current strength in the electrostatic system are obtained from the equation above. We have $[I] = \left[\frac{Q}{t}\right] = M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-2}$, the dimensions of current.

203. Ohm's Law.—In § 187 it was remarked that a body is distinguished as a good or a poor conductor by the rate at which it will equalize the potentials of two electrified conductors, if it be used to connect them. Manifestly this property of the substances forming a circuit, of conducting electricity rapidly or otherwise, will influence the strength of the current in the circuit. It was shown on theoretical considerations, in 1827, by Ohm of Berlin, that in a homogeneous conductor which is kept constant, the current varies directly with the difference of potential between the terminals. If R represent a factor, constant for each conductor, *Ohm's law* is expressed in its simplest form by

$$IR = V_1 - V_2. \tag{86}$$

The quantity R is called the *resistance* of the conductor. If the difference of potential be maintained constant, and the conductor be altered in any way that does not introduce an internal electromotive force, the current will vary with the changes in the conductor, and there will be a different value of R with each change in the conductor. The quantity R is therefore a function of the nature and materials of the conductor, and does not depend on the current or the difference of potential between the ends of the conductor. Since it is the ratio of the current to the difference of potential, and since we know these quantities in electrostatic units, we can measure R in electrostatic units. From the dimensions of I and $(V_1 - V_2)$ we may obtain the dimensions of R. They are in electrostatic units

$$[R] = \left[\frac{V_1 - V_2}{I}\right] = L^{-1}T.$$

To generalize Ohm's law for the whole circuit, let us consider a special circuit which may serve as a type. It shall consist of a voltaic cell containing acidulated water, in which are immersed a zinc and a platinum plate, joined together by a platinum wire outside the liquid (Fig. 62). Consider a point in the liquid just outside the zinc; if the potential of a point near it, just inside the zinc, be V_Z, then the potential at the point considered is $V_Z + Z/L$, if Z/L represent the sudden change in potential across the surface of separation. The potential at a point in the liquid just outside the platinum is V_L, and by the elementary form of Ohm's law already considered we have

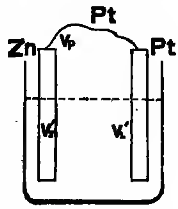

FIG. 62.

$$I = \frac{V_Z + Z/L - V_L}{R_L}.$$

In the same way the current in the platinum and platinum wire is expressed by

$$I = \frac{V_L + L/P - V_P}{R_P};$$

and in the zinc by

$$I = \frac{V_P + P/Z - V_Z}{R_Z}.$$

Now these currents are all equal, for there is no accumulation of electricity anywhere in the circuit. Hence

$$I = \frac{V_Z + Z/L - V_L}{R_L} = \frac{V_L + L/P - V_P}{R_P}$$
$$= \frac{V_P + P/Z - V_Z}{R_Z},$$

or

$$I = \frac{Z/L + L/P + P/Z}{R_L + R_P + R_Z}.$$

But the numerator is the sum of all the differences of potential in the circuit taken in one direction, or the measure of the electromotive force, and the denominator is the total resistance of the circuit. It may then be stated more generally as *Ohm's law* that in any circuit the current equals the electromotive force divided by the resistance, or

$$I = \frac{E}{R}. \tag{87}$$

204. Specific Conductivity and Specific Resistance. — If two points be kept at a constant difference of potential, and joined by a homogeneous conductor of uniform cross-section, it is found that the current in the conductor is directly proportional to its cross-section and inversely as its length. The current also depends upon the nature of the conductor. If conductors of similar dimensions, but of different materials, are used, the current in each is proportional to a quantity called the *specific conductivity* of the material. The numerical value of the current set up in a conducting cube, with edges of unit length, by unit difference of potential between two opposite faces, is the measure of the conductivity of the material of the cube. The reciprocal of this number is the *specific resistance* of the material. If ρ represent the specific resistance of the conducting material, S the cross-section and l the length of a portion of the conductor of uniform cross-section between two points at potentials V_1 and V_2, Ohm's law for this special case can be presented in the formula

$$I = \frac{S(V_1 - V_2)}{l\rho}. \tag{88}$$

The specific resistance is not perfectly constant for any one material, but varies with the temperature. In metals the specific resistance increases with rise in temperature; in liquids and in carbon it diminishes with rise in temperature. Upon this fact of change of resistance with temperature is based a very delicate instrument, called by Langley, its inventor, the *bolometer*, for the measurement of the intensity of radiant energy.

205. Joule's Law. — If we modify the equation $H = I(V_1 - V_2)$ by the help of Ohm's law, we obtain

$$H = I^2R. \tag{89}$$

The heat developed in a homogeneous portion of any circuit is equal to the square of the current in the circuit multiplied by the resistance of that portion. This relation was first experimentally proved by Joule in 1841, and is known after his name as *Joule's law*. It holds true for any homogeneous circuit or for all parts of a circuit which are homogeneous. The heat which is sometimes evolved by chemical action, or by the Peltier effect, occurs at non-homogeneous portions of the circuit.

206. Counter Electromotive Force in the Circuit.—In many cases the work done by the current does not appear wholly as heat developed in accordance with Joule's law.

Besides the production of heat throughout the circuit, work may be done during the passage of the current, in the decomposition of chemical compounds, in producing movements of magnetic bodies or other circuits in which currents are passing, or in heating junctions of dissimilar substances.

Before discussing these cases separately we will connect them all by a general law, which will at the same time present the various methods by which currents can be maintained. They differ from the simple case in which the work done appears wholly as heat throughout the circuit, in that the work done appears partly as energy available to generate currents in the circuit. To show this we will use the method given by Helmholtz and by Thomson. The total energy expended in the circuit in the time t, which is such that, during it, the current is constant, is IEt. It appears partly as heat, which equals I^2Rt by Joule's law, and partly as other work, which in every case is proportional to I, and can be set equal to IA, where A is a factor which varies with the particular work done. Then we have $IEt = I^2Rt + IA$, whence

$$I = \frac{E - \frac{A}{t}}{R}. \qquad (90)$$

It is evident from the equation that $E - \frac{A}{t}$ is an electromotive force, and that the original electromotive force of the circuit has been modified by the fact of work having been done by the current. In other words, the performance of the work IA in the time t by the circuit has set up a counter electromotive force $\frac{A}{t}$. The separated constituents of the chemical compound, the moved magnet, the heated junction, are all sources of electromotive force which oppose that of the original circuit. If then, in a circuit containing no impressed electromotive force, or in which $E = 0$, there be brought an arrangement of uncombined chemical substances which are capable of combination, or if in its presence a magnet or closed current be moved, or if a junction of two dissimilar parts of the circuit be heated, there will be set up an electromotive force $\frac{A}{t}$, and a current $I = \frac{\frac{A}{t}}{R}$. Any of these methods may then be used as the means of generating a current. The first gives the ordinary battery currents of Volta, the second the induced currents discovered by Faraday, and the third the thermo-electric currents of Seebeck.

CHAPTER IV.

CHEMICAL RELATIONS OF THE CURRENT.

207. Electrolysis.—It has been already mentioned that, in certain cases, the existence of an electrical current in a circuit is accompanied by the decomposition into their constituents of chemical compounds forming part of the circuit. This process, called *electrolysis*, must now be considered more fully. It is one of those treated generally in § 206, in which work other than heating the circuit is done by the current. That work is done by the decomposition of a body the constituents of which, if left to themselves, tend to recombine, is evident from the fact that, if they be allowed to recombine, the combination is always attended with the evolution of heat or the appearance of some other form of energy. The amount of heat developed, or the energy gained, is, of course, the measure of the energy lost by combination or necessary to decomposition.

A free motion of the molecules of a body, associated with close contiguity, seems to be necessary in order that it may be decomposed by the current. Only liquids, and solids in solution or fused, have been electrolysed. Bodies which can be decomposed were called by Faraday, to whom the nomenclature of this subject is due, *electrolytes*. The current is usually introduced into the electrolyte by solid terminals called *electrodes*. The one at the higher potential is called the positive electrode, or *anode;* the other, the negative electrode, or *cathode*. The two constituents into which the electrolyte is decomposed are called *ions*. One of them appears at the anode and is called the ***anion***, the other at the cathode and is called the ***cation***.

For the sake of clearness we will describe some typical cases of electrolysis. The original observation of the evolution of gas when the current was passed through a drop of water, made by Nicholson and Carlisle, was soon modified by Carlisle in a way which is still generally in use. Two platinum electrodes are immersed in water slightly acidulated with sulphuric acid, and tubes are arranged above them so that the gases evolved can be collected separately. When the current is passing, bubbles of gas appear on the electrodes. When they are collected and examined, the gas which appears at the anode is found to be oxygen, and that which appears at the cathode to be hydrogen. The quantities evolved are in the proportion to form water. This appears to be a simple decomposition of water into its constituents, but it is probable that the acid in the water is first decomposed, and that the constituents of water are evolved by a secondary chemical reaction.

An experiment performed by Davy, by which he discovered the elements potassium and sodium, is a good example of simple electrolysis. He fused caustic potash in a platinum dish, which was made the anode, and immersed in the fused mass a platinum wire as cathode. Oxygen was then evolved at the anode, and the metal potassium was deposited on the cathode. This is the type of a large series of decompositions.

If, in a solution of zinc sulphate, a plate of copper be made the anode and a plate of zinc the cathode, there will be zinc deposited on the cathode and copper taken from the anode, so that, after the process has continued for a time, the solution will contain a quantity of cupric sulphate. This is a case similar to the electrolysis of acidulated water, in which the simple decomposition of the electrolyte is modified by secondary chemical reaction.

If two copper electrodes be immersed in a solution of cupric sulphate, copper will be removed from the anode and deposited on the cathode, without any important change occurring in the character or concentration of the electrolyte. This is an example of the special case in which the secondary reactions in the electrolyte exactly balance the work done by the current in decomposition, so that on the whole no chemical work is done.

208. **Faraday's Laws.**—The researches of Faraday in electrolysis developed two laws, which are of great importance in the theory of chemistry as well as in electricity.

(1) The amount of an electrolyte decomposed is directly proportional to the quantity of electricity which passes through it; or, the rate at which a body is electrolysed is proportional to the current strength.

(2) If the same current be passed through different electrolytes, the quantity of each ion evolved is proportional to its chemical equivalent.

If we define an *electro-chemical equivalent* as the quantity of any ion which is evolved by unit current in unit time, then the two laws may be summed up by saying:

The number of electro-chemical equivalents evolved in a given time by the passage of any current through any electrolyte is equal to the number of units of electricity which pass through the electrolyte in the given time.

The electro-chemical equivalents of different ions are proportional to their chemical equivalents. Thus, if zinc sulphate, cupric sulphate, and argentic chloride be electrolysed by the same current, zinc is deposited on the cathode in the first case, copper in the second, and silver in the third. The amounts by weight deposited are in proportion to the chemical equivalents, 32.6 parts of zinc, 31.7 parts of copper, and 108 parts of silver.

209. **The Voltameter.**—These laws were used by Faraday

to establish a method of measuring current by reference to an arbitrary standard. The method employs a vessel containing an electrolyte in which suitable electrodes are immersed, so arranged that the products of electrolysis, if gaseous, can be collected and measured or, if solid, can be weighed. This arrangement is called a *voltameter*. If the current strength be desired, the current must be kept constant in the voltameter by suitable variation of the resistance in the circuit during the time in which electrolysis is going on.

Two forms of voltameter are in frequent use.

In the first form there is, on the whole, no chemical work done in the electrolytic process. The system consisting of two copper electrodes and cupric sulphate as the electrolyte is an example of such a voltameter. The weight of the copper deposited on the cathode measures the current.

The second form depends for its indications on the evolution of gas, the volume of which is measured. The water voltameter is a type, and is the form especially used. The gases evolved are either collected together, or the hydrogen alone is collected. The latter is preferable, because oxygen is more easily absorbed by water than hydrogen and an error is thus introduced when the oxygen is measured.

210. Measure of the Counter Electromotive Force of Decomposition.—In the general formula developed in § 206, the quantity IA represents the energy expended in the circuit which does not appear as heat developed in accordance with Joule's law. In the present case it is the energy expended during electrolysis in decomposing chemical compounds and in doing mechanical work. In many cases the mechanical work done is not appreciable; but when a liquid like water is decomposed into its constituent gases, work is done by the expansion of the gases from their volume as water to their volume as gases. Let e represent the electro-chemical equivalent of one of the ions, and θ the heat evolved by the combination of a unit mass of this ion with an equivalent mass of the other ion, in which is included the heat equivalent of the mechanical work done if the state of aggregation change. Then Ie will represent the number of electro-chemical equivalents evolved, and $Ie\theta$ will represent the energy expended, which appears as chemical separation and mechanical work. This is equal to IA; whence $A = e\theta$. All these quantities are measured in absolute units. The quantity $e\theta$ represents the energy required to separate the quantity e of the ion considered from the equivalent quantity of the other ion, and to bring both constituents to their normal condition.

If the electrolytic process go on uniformly for a time t, so that equal quantities of the ion considered are evolved in equal times, we have $\frac{A}{t} = \frac{e\theta}{t}$. Now, $\frac{A}{t}$ represents the counter-electromotive force set up in the circuit by electrolysis. Hence the electromotive force set up in the electrolytic process may be measured in terms of heat units; or, since these heat units are measures of chemical affinity, the same relation gives a measure of chemical affinity in terms of electromotive force.

It often is the case that the two ions which appear at the electrodes are not capable of direct recombination, as has been tacitly assumed in the definition of θ. A series of chemical exchanges is always possible, however, which will restore the ions as constituents of the electrolyte, and the total heat evolved for a unit mass of one ion during the process is the quantity θ.

The theory here presented is abundantly verified by the experiments of Joule, Favre and Silbermann, Wright and others.

211. Positive and Negative Ions.—Experiment shows that certain of the bodies which act as ions usually appear at the cathode, and certain others at the anode. The former are called *electro-positive* elements; the latter, *electro-negative* elements. Faraday divided all the ions into these two classes, and thought that every compound capable of electrolysis was made up of one electro-positive and one electro-negative ion. But the distinction is not absolute. Some ions are electro-positive in one combination and electro-negative in another. Berzelius made an attempt to arrange the ions in a series, such that any one ion should be electro-positive to all those above it and electro-negative to all those below it. It is questionable whether a rigorous arrangement of the ions is at the present time possible.

212. Theory of Electrolysis.—When any attempt is made to explain the behavior of the ions in the process of electrolysis, grave difficulties are met with at once. The foundation of all the present theories is found in the theory published by Grotthus in 1805. He considers the constituent ions of a molecule as oppositely electrified to an equal amount. When the current passes, owing to the electrical attractions of the electrodes, the molecules arrange themselves in lines with their similar ends in one direction, and then break up. The electro-negative ion of one molecule moves toward the positive electrode and meets the electro-positive ion of the neighboring molecule, with which it momentarily unites. At the ends of the line an electro-negative ion with its charge is freed at the anode, and an electro-positive ion with its charge at the cathode. This process is repeated indefinitely so long as the current passes.

Faraday modified this view, in that he ascribed the arrangement or *polarization* of the molecules, and their disruption, to the stress in the medium which was the cardinal point in his electrical theories. Otherwise he held closely to Grotthus' theory. He showed that the state of polarization existed in the electrolyte by means of fine silk threads immersed in it. These arranged themselves along the lines of electrical stress.

Other phenomena, however, show that Grotthus' hypothesis can only be treated as a rough mechanical illustration of the main facts.

Joule showed that during electrolysis there is a development of heat at the electrodes, in certain cases, which is not accounted for by the elementary theory above given. It must depend upon a more complicated process of electrolysis than the one we have described.

The results of researches on the so-called *wandering* of the ions are also at variance with Grotthus' theory. If the electrolysis of a copper salt, in a cell with a copper anode at the bottom, be examined, it will be found that the solution becomes more concentrated about the anode and more dilute about the cathode. These changes can be detected by the color of the parts of the solution, and substantiated by chemical analysis. If this result be explained by Grotthus' theory, the explanation furnishes at the same time a numerical relation between the ions which have wandered to their respective regions in the electrolyte which is not in accord with experiment.

Another peculiar phenomenon, known as *electrical endosmose*, may also be mentioned in this connection. It is found that, if the electrolyte be divided into two portions by a porous diaphragm, there is a transfer of the electrolyte toward the cathode, so that it stands at a higher level on the side of the diaphragm nearer the cathode than on the other. This fact was discovered by Reuss in 1807, and has been investigated by Wiedemann and Quincke. They found that the amount of the electrolyte transferred is proportional to the current strength, and independent of the extent of surface or the thickness of the diaphragm. Quincke has also demonstrated a flow of the electrolyte toward the cathode in a narrow tube, without the intervention of a diaphragm. Those electrolytes which are the poorest conductors show the phenomenon the best. In a very few cases the motion is towards the anode. The material of which the tube is composed influences the direction of flow. It has also been shown that solid particles move in the electrolyte, usually towards the anode.

To explain these phenomena, Quincke has brought forward a theory of electrolysis which is widely different from Grotthus' simple hypothesis, but is too complicated for presentation here.

It is an objection against Grotthus' theory, and indeed against Thomson's method given in § 210 of connecting chemical affinity and electromotive force, that, on those theories, it would require an electromotive force in the circuit greater than $\frac{A}{t}$, the counter electromotive force in the electrolytic cell, to set up a current, and that the current would begin suddenly, with a finite value, after this electromotive force was reached. On the contrary, experiments show that the smallest electromotive force will set up a current in an electrolyte and even maintain one constantly, though the current strength may be extremely small.

This is explained by Clausius by the help of the theory of the constitution of liquids which is now generally adopted. He conceives the molecules of the electrolyte to be moving about with different velocities. He thinks that occasionally the attraction between two opposite ions of two neighboring molecules may become greater than that between the constituents of the molecules. In that case the molecules are broken up, the two attracting ions combine to form a new molecule, and two opposite ions are set free. These may at once combine to form another new molecule, or they may wander through the mass until they meet with other ions, with which they can unite to again form molecules. He thinks that the electromotive force in the circuit, while not great enough to effect a decomposition of the electrolyte, may yet be sufficient to determine the direction of motion of these unpaired ions, so that they move, on the whole, towards their respective electrodes. Every theory of electrolysis assumes that the transfer of electricity is, in some way, connected with the transfer of the ions; hence on Clausius' theory there will be a current and an evolution of the ions with any electromotive force in the circuit, however low. This current would at once cease if the ions were to collect on the electrodes, and set up a permanent counter electromotive force; but the same reasoning as has just been used will show that the liberated ions, if not formed in such quantities as to collect and pass out of the liquid as in true electrolysis, will wander back into the liquid again. On this theory the number of free ions of either kind ought to be greater near the electrode to which they tend to move.

While Clausius' theory fully accounts for the behavior of the ions, it does not explain their relations to the electrical current. No satisfactory theory of the relations of electricity to the molecules of matter has as yet been given.

213. Voltaic Cells.—From the discussion given in § 206 it is obvious that, if an arrangement be made, in a circuit, of substances capable of uniting chemically and such as would result from electrolysis, there will result an electromotive force in such a sense as to oppose the current which would effect the electrolysis. If, then, the electrodes of an electrolytic cell in which this electromotive force exists be joined by a wire, a current will be set up through the wire in the opposite direction to the one which would continue the electrolysis, and the ions at the electrodes will recombine to form the electrolyte. There is thus formed an independent source of current, the *voltaic cell.* The electrode in connection with the electro-negative ion is called the *positive pole*, and that in connection with the electro-positive ion the *negative pole.*

Thus, if after the electrolysis of water in a voltameter, in which the gases are collected separately in tubes over platinum electrodes, the electrodes be joined by a wire, a current will be set up in it, and the gases will gradually, and at last totally, disappear, and the current will cease. The current which decomposes the water is conventionally said to flow through the liquid from the anode to the cathode, from the electrode above which oxygen is collected to the electrode above which hydrogen is collected. The current existing during the recombination of the gases flows through the liquid from the hydrogen electrode to the oxygen electrode, or outside the liquid from the positive to the negative pole. Such an arrangement as is here described was devised by Grove, and is called the *Grove's gas battery*.

A combination known as *Smee's cell* consists of a plate of zinc and one of platinum, immersed in dilute sulphuric acid. It is such a cell as would be formed by the complete electrolysis of a solution of zinc sulphate, if the zinc plate were made the cathode. When the zinc and platinum plates are joined by a wire, a current is set up from the platinum to the zinc outside the liquid, and the zinc combines with the acid to form zinc sulphate. The hydrogen thus liberated appears at the platinum plate, where, since the oxygen which was the electro-negative ion of the hypothetical electrolysis by which the cell was formed does not exist there ready to combine with it, it collects in bubbles and passes up through the liquid. The presence of this hydrogen at once lowers the current from the cell, for it sets up a counter electromotive force, and also diminishes the surface of the platinum plate in contact with the liquid, and thus increases the resistance of the cell. It may be partially removed by mechanical movements of the plate or by roughening its surface. The counter electromotive force is called the *electromotive force of polarization*. It occurs soon after the circuit is joined up in all cells in which only a single liquid is used, and very much diminishes the currents which are at first produced.

Advantage is taken of secondary chemical reactions to avoid this electromotive force of polarization. The best example, and a cell which is of great practical value for its cheapness, durability, and constancy, is the *Daniell's cell*. Two liquids are used, solutions of cupric sulphate and zinc sulphate. They are best separated from one another by a porous porcelain diaphragm. A plate of copper is immersed in the cupric sulphate, and a plate of zinc in the zinc sulphate. The copper is the positive pole, the zinc the negative pole. When the circuit is made and the current passes, zinc is dissolved, the quantity of zinc sulphate increases and that of the cupric sulphate decreases, and copper is deposited on the copper plate. To prevent the destruction of the cell by the consumption of the cupric sulphate, crystals of the salt are placed in the solution. The electromotive force of this cell is evidently due to the loss of energy in the substitution of zinc for copper in the solution of cupric sulphate. It may be calculated by the formula of § 210. The experiments of Kohlrausch give for zinc in C. G. S. units, $e = 0.003411$, where the system of units employed is the electromagnetic (§ 218). Favre and Silbermann give for θ, in the chemical process here involved, 714 gram-degrees or lesser calories. The mechanical equivalent of one gram-degree is 41,595,000. Hence we obtain for the electromotive force of a Daniell's cell in C. G. S. electromagnetic units the value $1.013 \cdot 10^8$. The value as found by direct experiment is about $1.1 \cdot 10^8$ in C. G. S. electromagnetic units.

There are many other forms of cell, which are all valuable for certain purposes. One of the best known is the *Grove's cell.* It has for positive pole a platinum plate, immersed in strong nitric acid, and for negative pole a zinc plate, immersed in dilute sulphuric acid. The two liquids are separated by a porous porcelain diaphragm. When the current passes, the zinc is dissolved. The hydrogen freed is oxidized by the nitric acid, which is gradually broken up into other compounds. The electromotive force of the Grove's cell is very high, being about $1.95 \cdot 10^8$ C. G. S. electromagnetic units.

The *secondary cell of Planté* is an example of a cell made directly by electrolysis, as has been assumed in the preliminary discussion. The electrodes are both lead plates, and the electrolyte dilute sulphuric acid. When a current is passed through the cell, the oxygen evolved on the anode combines with the lead to form peroxide of lead, which coats the surface of the electrode. When the cell is inserted in a circuit, a current is set up, the peroxide is reduced to a lower oxide, and the metallic lead of the other plate is oxidized.

The *Latimer-Clarke standard cell* is of great value as a standard of electromotive force. As it polarizes at once if a current pass through it, it should never be joined up in a closed circuit. The positive pole consists of pure mercury, which is covered by a paste made by boiling mercurous sulphate in a saturated solution of zinc sulphate. The negative pole consists of pure zinc resting on the paste. Contact with the mercury is made by means of a platinum wire. As no gases are generated, this cell may be hermetically sealed against atmospheric influences. According to the measurements of Rayleigh, the electromotive force of this cell is very constantly $1.435 \cdot 10^8$ C. G. S. electromagnetic units at $15°$ Cent.

214. Theories of the Electromotive Force of the Voltaic Cell.—The plan followed in the preceding discussions has rendered it unnecessary for us to adopt any theory to explain the cause of the electromotive force of the voltaic cell. The different theories which have been advanced may be classed under one of two general theories, the contact theory and the chemical theory. On the *contact theory*, as advanced by Volta and supported by Thomson and others, the difference of potential which exists between two heterogeneous substances in contact is due to molecular interactions across the surface of contact, or, as it is commonly stated, is due merely to the contact. The *chemical theory*, as advocated by Faraday and Schönbein, holds that the difference of potential considered cannot arise unless chemical action or a tendency to chemical action exist at the surface of contact.

Numerous experiments have shown that the sum of all the differences of potential at the surfaces of contact of the various substances making up any voltaic cell is equal to the electromotive force of that cell. This is true even when the cell is formed solely of liquid elements. On the contact theory this electromotive force is due merely to the several contacts, while the chemical actions of the cell begin only when the circuit is made, and supply the energy for the maintenance of the current. The quantity of heat produced at a junction of dissimilar substances by the passage of a current (§ 233) is such as to show, however, that the differences of potential thus measured are not the true differences of potential due to the contact of the substances tested, but must depend in part upon the action of the air or other medium by which these substances are surrounded. The supporters of the chemical theory point to this fact as evidence that the chemical action of the medium is concerned in the production of the difference of potential observed.

On either theory it is clear that the energy maintaining the current must have its origin in the chemical actions which go on in the voltaic cell.

215. Capillary Electrometer.—It has been stated that a difference of potential exists between a metal and a fluid electrolyte in contact with it. There will then exist on the surfaces of the metal and the electrolyte in contact with it such an electrical distribution as exists in a charged condenser of which the plates are very near together.

One arrangement by which the effects due to this distribution may be observed was devised by Lippmann. It consists of a vertical glass tube, drawn out at its lower end in a capillary tube. The capillary tube dips into dilute sulphuric acid, which rests on mercury in the bottom of the vessel containing it. Mercury is poured into the vertical tube until its pressure is such that the capillary portion of the tube is nearly filled with it. When the mercury in the vessel is joined with the positive pole of a voltaic cell, and that in the tube with the negative pole, the meniscus in the capillary tube moves upward, in the sense in which it would move if its surface tension were increased. This movement may best be explained by the help of the theory of electrolysis given by Clausius (§ 212). So long as there exists an electromotive force in the circuit, positive and negative ions will be released on their respective electrodes. If we assume that they are associated with the transfer of electricity in the circuit in such a way that it is transferred from them to the electrodes, such a movement of the ions would give rise to a modification of the distribution on the surfaces of contact. In the case now under consideration the charge on the meniscus is in part neutralized by the charge transferred with one of the ions. The true surface tension of the surface of separation between the mercury and the liquid is, on this theory, lessened by the presence of the electrical charge on the surfaces of contact, owing to the interaction of the parts of the charge in a manner similar to that described in § 191. If, therefore, any diminution of this charge occur, a seeming increase of the surface tension will be observed. On this theory the true surface tension of the surface of separation is the value observed when the mercury and liquid are at the same potentials, and this value is a maximum. The experiments of A. König and Helmholtz show that such a maximum value exists in a manner consistent with the theory.

The arrangement described can manifestly be used to produce the effects just discussed only when the electromotive force introduced into the circuit is less than that required to cause active decomposition of the electrolyte. If any suitable electromotive force be introduced into the circuit, the theory here given assumes that the transfer of the ions goes on until the differences of potential on the surfaces of contact are such as to counterbalance the introduced electromotive force. The mercury column then comes to rest.

Lippmann constructed an apparatus similar to the one described, with the addition of an arrangement by which pressure can be applied to force the end of the mercury column in the capillary tube back to the fixed position which it occupies when no electromotive force is introduced into the circuit. He found that when small electromotive forces were introduced, the pressures required to bring the end of the column back to the fixed position were proportional to the electromotive forces. He hence called this apparatus a *capillary electrometer.*

Lippmann also found that if the area of the surface of separation between the mercury and the liquid in the capillary tube were altered by increasing the pressure and driving the mercury down the tube, a current was set up in a galvanometer inserted in the circuit, in a sense opposite to that which would change the area of the meniscus back to its original amount.

CHAPTER V.

MAGNETIC RELATIONS OF THE CURRENT.

216. Biot's Law.—Very soon after the discovery by Oersted of the fact that a magnet was acted upon by an electrical current brought near it, Biot and Savart instituted a series of experiments to determine the law of the force between a magnet and a current. They suspended a short magnet by a silk fibre, and so modified the earth's magnetic field near it, by means of magnets, that the suspended magnet pointed in any azimuth with equal freedom. A current was then passed through a long vertical wire near the magnet. It was observed that the magnet placed itself so that its poles were equally distant from the wire. The directions of the north pole and of the current were related as the motions of rotation and of propulsion in a right-handed screw. Then the magnet was set in oscillation, and the times of oscillation determined when the current was at different distances from the magnet, and when different currents were set up in the wire. From the first observation it follows that the force exerted between a magnet pole and a current is normal to the plane passing through the

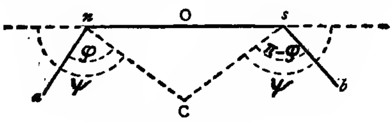

FIG. 63.

current and the magnet pole. For, suppose the current rising vertically out of the paper at C (Fig. 63), and suppose that it acts on the north pole of the magnet *ns* with a force represented by *na*, making any angle ϕ with the line *nC*. It is assumed as probable that the force on a south pole placed at *n* would be oppositely directed to *na*. The angle which the force *sb*, acting on the south pole *s*, makes with the line *sC* will then be $\pi - \phi$. Now the magnet is in equilibrium, hence the moments of the components of these forces at right angles to the magnet must be equal. The components are respectively $na \sin(\psi - \phi)$ and $sb \sin(\psi - (\pi - \phi))$. The lever arms *on* and *os* are equal, and it is assumed that, since the poles are at equal distances from the current, the forces *na* and *sb* are equal; therefore $\sin(\psi - \phi)$ must equal $\sin(\psi - (\pi - \phi))$, and this is true only when $\phi = \frac{\pi}{2}$. The lines of magnetic force about an infinite straight current are therefore circles, and the equipotential surfaces determined by these lines are planes passing through the current.

From the times of oscillation observed, it was proved that the force exerted is proportional directly to the strength of the magnet pole and to the strength of the current, and inversely to the distance between the pole and the current. Biot hence deduced a law for the action of each element of length of the current upon a magnet pole, which is expressed in the formula

$$f = \frac{mi \sin \alpha ds}{r^2}. \tag{91}$$

In this formula *m* represents the strength of the magnet pole, *i* the current strength measured in electromagnetic units, *ds* the element of the current, *r* the distance between that element and the magnet pole, and α the angle between *r* and *ds*. It is easy to show that the force exerted by a long straight current, observed by Biot to be inversely as the distance from the current, is consistent with this law. For simplicity we will consider an infinitely long straight current. Let the magnet pole m be at the point P (Fig. 64). Let QR be the current

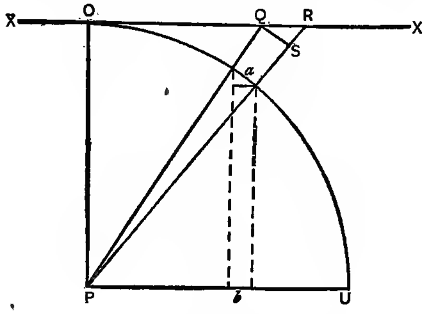

FIG. 64.

element ds, and PO the perpendicular distance between the pole and the current. Then Biot's law gives for the force exerted by the element QR the expression $\frac{miQR}{PR^2} \cdot \frac{PO}{PR}$. In the limit, as QR becomes indefinitely small, the triangles QRS and POR become similar. Hence QS equals $\frac{QR.PO}{PR}$, and the expression for the force becomes $\frac{miQS}{PR^2}$. If about P, with PO as radius, we draw the arc OU, the elementary arc a in the limit equals $\frac{QS.PO}{PR}$, and the projection b of the arc a on the line PU equals $\frac{aPO}{PR}$. Using these values, the expression $\frac{miQS}{PR^2}$ becomes $\frac{mib}{PO^2}$. There will be a similar expression for the force due to any other element. The total force due to the whole current will be equal to the constant factor $\frac{mi}{PO^2}$ multiplied by the sum of all the projections corresponding to b. This sum, for the infinite current, is manifestly $2PU = 2PO$. Hence the total force is $\frac{2mi}{PO}$; or, it is inversely as the distance PO between the pole and the current.

217. **Equivalence of a Closed Circuit and a Magnetic Shell.**—The law of the force between a pole and a current, which has been stated, leads to the conclusion that a very small closed plane circuit, carrying a current, will act upon a magnet pole at a distance from it in the same way as a magnetic shell, of which the edge coincides with the contour of the circuit, and the strength equals the strength of the current. To show this we will use a rectangular circuit with indefinitely small sides. We will place the origin (Fig. 65) at the centre of the rectangle, and draw the x axis perpendicular to the plane of the rectangle, and the y and z axes parallel with its sides. For convenience, we will call the length of the sides parallel to the y axis $2s$, and of those parallel to the z axis $2s'$.

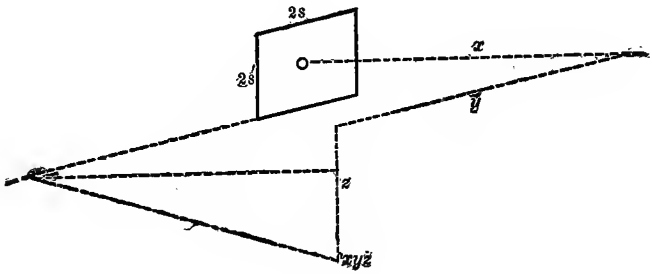

FIG. 65.

We assume that a current of strength i traverses the boundary of the rectangle in a direction related to the positive direction of the x axis, as the motions of rotation and propulsion are related in a right-handed screw.

If the magnet pole be at the point (xyz), the force on it due to one side, $2s$, is, as stated in Eq. (90), proportional to the length $2s$, is inversely as the square of the distance

$$(x^2 + y^2 + (z - s')^2)^{\frac{1}{2}},$$

and is proportional to the sine of the angle between the line joining (xyz) and the element $2s$ and the direction of that element. This sine is expressed by $\frac{(x^2 + (z - s')^2)^{\frac{1}{2}}}{(x^2 + y^2 + (z - s')^2)^{\frac{1}{2}}}$. The total force due to the element is then $\frac{mi2s(x^2 + (z - s')^2)^{\frac{1}{2}}}{(x^2 + y^2 + (z - s')^2)^{\frac{3}{2}}}$. This force is at right angles to the plane passing through the direction of the element $2s$ and the perpendicular from (xyz) on the direction of that element. We shall investigate in turn the components along the three axes. That along the x axis is found by multiplying the total force by $\frac{z - s'}{(x^2 + (z - s')^2)^{\frac{1}{2}}}$. The expression for the component along the x axis then becomes

$$\frac{2mis(z - s')}{(x^2 + y^2 + (z - s')^2)^{\frac{3}{2}}}.$$

We will expand $(z - s')^2$ in the denominator, reject the term s'^2, remembering that the sides are indefinitely small, and write for brevity $x^2 + y^2 + z^2 = r^2$. We then have this component expressed by $\frac{2mis(z - s')}{(r^2 - 2zs')^{\frac{3}{2}}}$. Similar expressions hold for the components due to each of the other sides, with the difference that those due to opposite sides must have different signs. We call those positive which are directed along the positive direction of x.

We will write the four components, and opposite them their expansions in ascending powers of s or s', rejecting all terms containing the second or higher powers of s.

$$-\frac{2mis(z-s')}{(r^2-2zs')^{\frac{3}{2}}} = -2mis(z-s')(r^{-3}+3s'zr^{-5});$$
$$+\frac{2mis(z+s')}{(r^2+2zs')^{\frac{3}{2}}} = +2mis(z+s')(r^{-3}-3s'zr^{-5});$$
$$-\frac{2mis'(y-s)}{(r^2-2ys)^{\frac{3}{2}}} = -2mis'(y-s)(r^{-3}+3syr^{-5});$$
$$+\frac{2mis'(y+s)}{(r^2+2ys)^{\frac{3}{2}}} = +2mis'(y+s)(r^{-3}-3syr^{-5}).$$

If we write out the sum of these expressions, rejecting all terms of the dimensions of s^3, we obtain as the component along the x axis of the force due to the whole circuit the expression $-4miss'\left(\frac{3y^2+3z^2}{r^5}-\frac{2}{r^3}\right)$. The term in parenthesis can be written $\frac{3y^2+3z^2-2r^2}{r^5}=\frac{r^2-3x^2}{r^5}$. The factor $4ss'$ is equal to a, the area of the rectangle. The force along the x axis is then finally

$$-mia\left(\frac{1}{r^3}-\frac{3x^2}{r^5}\right). \qquad (92)$$

For the component along the y axis we have to consider only the forces due to the sides $2s'$, for the other sides have no tendency to move the pole parallel with themselves. The components of these forces along y, that one being called positive which is in the positive direction of y, are $+2mis'\frac{x}{(r^2-2ys)^{\frac{3}{2}}}$ and $-2mis'\frac{x}{(r^2+2ys)^{\frac{3}{2}}}$. The sum of these components is

$$4miss'\frac{3xy}{r^5}=mia\frac{3xy}{r^5}. \qquad (92a)$$

Similarly the total component along the z axis is

$$mia\frac{3xz}{r^5}. \tag{92b}$$

Now to compare these forces with those due to a magnetic shell of the indefinitely small area a and strength j, we use the result of the discussion in § 182, that the potential of such a shell at any external point is $j\omega$. In that discussion the convention was made that the positive face of the shell was turned toward the positive direction of the x axis. We then have ω, the solid angle subtended by the shell as seen from the point P, equal to

$$\frac{a\cos\theta}{r^2} = \frac{ax}{r^3} = a\frac{x}{(x^2+y^2+z^2)^{\frac{3}{2}}}.$$

The potential at the point P is then

$$V = ja\frac{x}{(x^2+y^2+z^2)^{\frac{3}{2}}}.$$

To find the forces along the three axes we must find the rate of change of this potential with respect to space. To do this for the x axis, let x increase by a small increment Δx; then the potential will take a small increment ΔV. We will have

$$V + \Delta V = ja\frac{x+\Delta x}{((x+\Delta x)^2+y^2+z^2)^{\frac{3}{2}}};$$

and as Δx becomes indefinitely small,

$$V + \Delta V = ja\frac{x+\Delta x}{(r^2+2x\Delta x)^{\frac{3}{2}}}.$$

Expanding this expression, rejecting all terms containing the second or higher powers of Δx, we obtain

$$V + \Delta V = ja\left(\frac{x}{r^3} - \frac{3x^2 \Delta x}{r^5} + \frac{\Delta x}{r^3}\right).$$

From this we have further

$$\frac{\Delta V}{\Delta x} = ja\left(\frac{1}{r^3} - \frac{3x^2}{r^5}\right).$$

In the limit, as Δx becomes indefinitely small, this is the rate of change of potential along the x axis at the point (xyz).

The force along the x axis on a unit magnet pole at the point (xyz) is this rate of change of potential taken with the opposite sign. Hence the force on the magnet pole m at that point is $-mja\left(\frac{1}{r^3} - \frac{3x^2}{r^5}\right)$. Similarly the forces along the y and z axes can be found to be respectively $mja\frac{3xy}{r^5}$ and $mja\frac{3xz}{r^5}$.

If these expressions be compared with the expressions for the components of force arising from the action of the rectangular current, they will be seen to be completely identical, provided that the unit of current be so selected that the factors i and j are equal.

If the current in the circuit be reversed, the components of force due to it remain the same in amount but are opposite in direction. The direction of current in the circuit which will render its action completely identical with that of the magnetic shell may be readily stated. Let us draw a line through the magnetic shell, tangent to the lines of force, from the negative to the positive face, and call its direction the positive direction of the lines of force. Then the current in the equivalent circuit is such that its direction is related to the positive direction of the lines of force as the motions of rotation and propulsion are related in a right-handed screw.

It may now be shown that a finite circuit of any form carrying a current i is equivalent to a magnetic shell of uniform strength j, the edge of which coincides with the circuit. For a finite circuit may be conceived to be made up of an assemblage of elementary circuits of the kind considered, lying contiguous to one another in the surface bounded by the contour of the circuit. Everywhere the currents of one of these elementary circuits is neutralized by the equal and infinitely near currents in the opposite direction of the contiguous circuits, except at the boundary, where all the elementary currents are in the same direction and are equivalent to the current in the circuit. This reasoning will be plain at once from Fig. 66. The forces due to such a current will then be equal to the forces due to a magnetic shell made up of elements which correspond to the elementary circuits. The systems of lines of force due to the shell and the equivalent circuit will be precisely similar in form and distribution. They will differ, however, in this, that the line of force joining two contiguous points on opposite faces of the shell will be interrupted by the shell, while in the case of the circuit it passes through the circuit as a continuous line enclosing the current. If a unit positive magnet pole were placed at a point on the positive face of a magnetic shell, it would move along a line of force to a point infinitely near the one from which it started, but on the opposite or negative face of the shell, and during the movement it would do an amount of work expressed by $4\pi j$. This same amount of work would be done upon it if it were brought back by any path to the point from which it started, so that the total work done in the

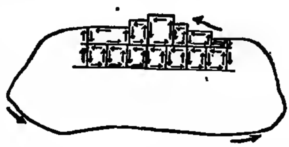

FIG. 66.

closed path is zero. If, on the other hand, the pole were moving under the influence of the circuit equivalent to the magnetic shell, it would move, as in the case of the shell, along the line of force from the positive to the negative face of the circuit, and in so doing would do work equal to $4\pi i$. But from the fact that the line of force on which it is moving is continuous, and that the force in the field is everywhere finite, it would pass over the infinitesimal distance between the point on the negative face and the one on the positive face, from which it started, without doing any finite work. The system would then have returned to its original condition, and work equal to $4\pi i$ would have been done. This is expressed by saying that the potential of a closed current is *multiply-valued.* The work done during any movement depends not only on the position of the initial and final points in the path, as in the case of the ordinary single-valued gravitational, electrical, and magnetic potentials, but also on the path traversed by the moving magnet pole. Every time the path encloses the current, work equal to $4\pi i$ is done. The work done in moving by a path which does not enclose the current, from a point where the solid angle subtended by the circuit is $\omega_{,}$ to one where it is ω, is, as in the case of the magnetic shell, equal to $i(\omega_{,} - \omega)$. If the path further enclose the current n times, the work done is $4\pi ni$, so that the total work done, or the total difference of potential between the two points, is

$$V_{,} - V = i(\omega_{,} - \omega + 4\pi n), \tag{93}$$

where n may have any value from 0 to infinity.

The fact that the potential of a current is multiply-valued is well illustrated by any one of a series of experiments due to Faraday. If we imagine a wire frame forming three sides of a rectangle to be mounted on a support so as to turn freely about one of its sides as a vertical axis, while the free end of the opposite side dips in mercury contained in a circular trough of which the axis of rotation passes through the centre, and if we suppose a current to be sent through the axis and the frame, passing out through the mercury; then if a magnet be placed vertically with its centre on the level of the trough, and with either pole confronting the frame, the frame will rotate continuously about the axis.

Other arrangements are made by which more complicated rotations of circuits can be effected. If the circuit be fixed and the magnet movable, similar arrangements will give rise to motions of the magnet or to rotations about its own axis.

218. Electromagnetic Unit of Current.—The relation which has been discussed between a circuit and the equivalent magnetic shell affords a means of defining a unit of current different from that before defined in the electrostatic system. That current is defined as the *unit current*, which will set up the same magnetic field as that due to a magnetic shell of which the edge coincides with the circuit, and the strength is unity.

This definition is equivalent to the following one, which is sometimes given. If the force between a unit magnet pole and a current flowing in a plane circuit of unit length, every part of which is at unit distance from the pole, be the unit force, then the current is the unit current.

The equivalence of the two definitions may be proved as follows. Conceive a circular plane magnetic shell of strength j to be set up normal to the x axis, with its centre at the origin. Then at the centre $x = 0$, and from § 217 the component of force along the x axis due to each element of the shell is numerically $\frac{mja}{r^3}$. Now divide the circle up into any very great number of circular rings by striking circles about its centre with radii differing by the small distance d. The elementary areas making up any one of these rings are all at the same distance from the centre, and the force along the x axis due to the whole circle can be found by summing the areas of the rings divided by the cubes of their respective radii. Select the ring the radius of which is one half that of the circle, and call that radius l. The radii of the rings distant nd from this middle one are $l + nd$ and $l - nd$, and the areas of these rings are $2\pi d(l + nd)$ and $2\pi d(l - nd)$. The forces due to them are $\frac{2\pi mjd}{(l + nd)^2}$ and $\frac{2\pi mjd}{(l - nd)^2}$. These expressions are equal to $2\pi mj\left(\frac{d}{l^2} - \frac{2nd^2}{l^3}\right)$ and $2\pi mj\left(\frac{d}{l^2} + \frac{2nd^2}{l^3}\right)$ if we neglect the higher powers of d. The force due to both rings is $\frac{4\pi mjd}{l^2}$. As there are $\frac{l}{d}$ such pairs, the force due to the whole shell is $\frac{4\pi mj}{l}$. The force due to a circuit equivalent to the magnetic shell is $\frac{4\pi mi}{l}$. Since $l = \frac{R}{2}$, where R is the radius of the circle, the force along the x axis equals $\frac{2\pi im}{R}$.

If we adopt the second definition of unit current, and use Biot's formula for the action of a current on a magnet pole, the force due to a circular current, made up of current elements of length s, upon a pole at its centre is $\Sigma \frac{mis}{R^2}$. The sum of all the elements of the circle is $2\pi R$. Hence the force on this definition is also $\frac{2\pi im}{R}$.

The unit based upon these definitions is called the *electromagnetic* unit of current. It is fundamental in the construction of the electromagnetic system of units, in just the same way as the unit of quantity is fundamental in the electrostatic system.

In practice another *unit of current*, called the *ampere*, is used. It is equal to 10^{-1} C. G. S. electromagnetic units. The dimensions of the electromagnetic unit of current are those of the strength of a magnetic shell, or $[i] = M^{\frac{1}{2}}L^{\frac{1}{2}}T^{-1}$.

219. Lines of Magnetic Force.—It is convenient, in much of the discussion of the action of currents, to use the notion of lines of force, and to measure the strength of field, as explained in § 21, by the number of lines of force. For example, we may conceive the field about a magnet pole to be filled with conical tubes of force, of an angular aperture which is very small, and equal for all the cones, but otherwise entirely arbitrary. It is commonly assumed that each one of these cones represents a line of force. Then the solid angle subtended by any magnetic shell in the field, which is measured by the number of the cones contained in that solid angle, can be replaced by the number of lines of force which the boundary of the shell encloses.

If the magnet pole be free to move, it will move from a point of higher to a point of lower potential; that is, it will move in general to a point as near as possible to the negative face of the shell. If we make the convention that a line of force passes through a shell in the positive direction when it passes from the negative to the positive face, we may describe this motion as one of which the result is, that as many lines of force as possible pass through the shell in the positive direction. If the magnet pole be fixed, and the shell free to move, it follows, from the equality of action and reaction, that the shell will set itself so that as many lines of force as possible will pass through it in the positive direction. When the shell is not perfectly free to move, and in certain other special cases, it is sometimes convenient to use an equivalent statement, that the shell will move so that as few lines of force as possible pass through it in the negative direction.

These last conclusions are independent of the particular character of the magnetic field in which the shell is situated. It may then be stated generally, as a law governing the motions of magnetic shells or their equivalent electrical circuits in a magnetic field, that they tend to move so that as many lines of force as possible will pass through them in the positive direction. From the discussion in § 217 it may be seen that the positive direction of a line of force due to a current is related to the direction of the current in the circuit as the directions of propulsion and of rotation in a right-handed screw. To one looking at the negative face of a magnetic shell, the current in the equivalent circuit will travel with the hands of a watch.

If a part only of the closed circuit be free to move, it may be considered by itself as a magnetic shell, and it will move in accordance with the same law. We can therefore use this law to investigate the movements of circuits or parts of circuits due to the magnetic field in which they are placed.

220. Mutual Action of Two Currents.—In general, two plane circuits, if they be free to move, will so place themselves that the lines of force from the positive face of one will pass through the other in the positive direction, or through its negative face. The currents in the two circuits will then have the same direction. If they be placed so that unlike faces are opposed, they will move towards one another; if so that similar faces are opposed, they will move away from one another. Since in the first case the currents are in the same direction, and in the second in opposite directions, the law may be stated in another form: that circuits carrying currents in the same direction attract one another; in opposite directions, repel one another.

Parts of the circuits, if movable, follow the same law. For example, consider a circuit in the form of a wire square, free to turn about a vertical line passing through the centres of two opposite sides. If now a vertical wire, forming part of another circuit, be brought near one of the vertical sides of the square, that side will move towards the vertical wire, or away from it, according as the currents in the two wires are in the same or in opposite directions. It is clear that the maximum number of lines of force due to the fixed circuit pass through the movable circuit in the positive direction, when the two parallel portions carrying currents in the same direction are as near one another as possible; and that as few lines of force as possible pass through the movable circuit in the negative direction, when the two parallel portions carrying currents in opposite directions are as far from one another as possible.

221. Ampère's Law for the Mutual Action of Currents. —The laws of the action between electrical currents were first investigated by Ampère from a different point of view. From a series of ingenious experiments he deduced a law which expresses the action of a current element on any other current element. The action of any circuit on any other can be obtained from this law by summing the effects of all the elements. The complete deduction of the law from the experimental facts is too complicated to be given, but the experiments themselves are of great interest.

Ampère's method consisted in submitting a movable circuit or part of a circuit carrying a current to the action of a fixed circuit, and so disposing the parts of the fixed circuit that the forces arising from different parts exactly annulled one another, so that the movable circuit did not move when the current in the fixed circuit was made or broken. In the first two of his experiments the movable circuit consisted of a wire frame of the form shown in Fig. 67. The current passes into the frame by the points *a* and *b*, upon which the frame is supported. It is evident that the two halves of the frame tend to face in opposite directions in the earth's mag-

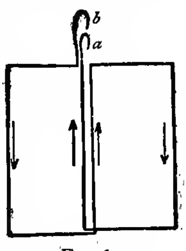

FIG. 67.

netic field, so that there is no tendency of the frame as a whole to face in any one direction rather than any other. If a long straight wire be placed near to one of the extreme vertical sides of the frame and a current be sent through it, that side will move towards the wire if the currents in it and in the wire be in the same direction, and will move away from the wire if the currents be in opposite directions.

If now this wire be doubled on itself, so that near the frame there are two equal currents occupying practically the same position, but in opposite directions, then no motion of the frame can be observed when a current is set up in the wire. This is Ampère's *first case* of equilibrium. It shows that the forces due to two currents, identical in strength and in position, but opposite in direction, are equal and opposite.

If the portion of the wire which is doubled back be not left straight, but bent into any sinuosities, provided these be small compared with the distance between the wire and the frame, still no motion of the frame occurs when a current is set up in the wire. This is Ampère's *second case* of equilibrium. It shows that the action of the elements of the curved conductor is the same as that of their projections on the straight conductor.

To obtain the *third case* of equilibrium, a wire, bent in the arc of a circle, is arranged so that it may turn freely about a vertical axis passing through the centre of the circle of which the wire forms an arc, and normal to the plane of that circle. The wire is then free to move only in the circumference of that circle, or in the direction of its own length. Two vessels filled with mercury, so that the mercury stands above the level of their sides, are brought under the wire arc, and raised until conducting contact is made between the wire and the mercury in both vessels. A current is then passed through the movable wire through the mercury. Then if any closed circuit whatever, or any magnet, be brought near the wire, it is found that the wire remains stationary. The deduction from this observation is that no closed circuit tends to displace an element of current in the direction of its length.

In the fourth experiment three circuits are used, which we may call respectively A, B, and C. They are alike in form, and the dimensions of B are mean proportionals to the corresponding dimensions of A and C. B is suspended so as to be free to move, and A and C are placed on opposite sides of B, so that the ratio of their distances from B is the same as the ratio of the dimensions of A to those of B. If then the same current be sent through A and C, and any current whatever through B, it is found that B does not move. The opposing forces due to the actions of A and C upon B are in equilibrium. From this *fourth case* of equilibrium is deduced the law that the force between two current elements is inversely as the square of the distance between them.

Ampère made the assumption that the action between two current elements is in the line joining them. From the four cases of equilibrium he then deduced an expression for the attraction between two current elements. It is

$$\frac{ii'\,ds\,ds'}{r^2}\left(2\cos\epsilon - 3\cos\theta\cos\theta'\right). \tag{94}$$

In this formula ds and ds' represent the elements of the two circuits, i and i' the strength of current in those circuits measured in electro-magnetic units, r the distance between the current elements, ϵ the angles made by the two elements with one another, θ and θ' the angles made by ds and ds' with r or r produced, the direction of the two elements being taken in the sense of their respective currents.

A remarkable result of this equation is that two current elements of the same circuit in the same straight line repel one another. The angle ϵ becomes $= 0$, and $\theta = \theta' = 0$; therefore the force given by the equation is $-\dfrac{ii'\,ds\,ds'}{r^2}$. Since this is negative it expresses a repulsion.

222. Solenoids and Electromagnets. — Ampère also showed that the action between two small plane circuits is the same as that between two small magnetic shells, and that a circuit, or system of circuits, may be constructed which is the complete equivalent of any magnet. A long bar magnet may be looked on as made up of a great number of equal and similar magnetic shells arranged perpendicular to the axis of the magnet, with their similar faces all in one direction. In order to produce the equivalent of this arrangement with the circuit, a long insulated wire is wound into a close spiral, straight and of uniform cross-section. The end of the wire is passed back through the spiral. When the current passes, the action of each turn of the spiral may be resolved into two parts, that due to the projection of the spiral on the plane normal to the axis, and that due to its projection on the axis. This latter component, for every turn, is neutralized by the current in the returning wire, and the action of the spiral is reduced to that of a number of similar plane circuits perpendicular to its axis. Such an arrangement is called a *solenoid.* The poles of a solenoid of very small cross-section are situated at its ends, and it is equivalent to a bar magnet uniformly magnetized.

If a bar of soft iron be introduced into the magnetic field within a solenoid it will become magnetized by induction. This combination is called an *electromagnet.* Since the strength of the magnetic field varies with the strength of the current in the solenoid, and with the number of layers of wire wrapped around the iron core, the magnetization of any bar of iron whatever may be raised to its maximum by increasing the current or the number of turns of wire.

223. Ampère's Theory of Magnetism. — Ampère based upon these facts a famous theory of magnetism which bears his name. He assumed that around every molecule of iron there circulates an electrical current, and that to such molecular currents are due all magnetic phenomena. He made no hypothesis with regard to the origin or the permanency of these currents. The theory agrees with Weber's hypothesis that magnetization consists in an arrangement of magnetic molecules. If we further adopt Thomson's explanation of the diamagnetic phenomena (§ 184), we may extend Ampère's theory to all matter, and assume that an electrical current circulates about every molecule. In order to account for the different magnetic susceptibilities of different bodies, it must also be assumed that these molecular currents are of different intensities in different kinds of matter.

Ampère's theory, however, admits another explanation of diamagnetism, which was given by Weber. He assumes that all diamagnetic molecules are capable of carrying molecular currents, but that those currents, under ordinary conditions, do not exist in them. When, however, a diamagnetic body is moved up to a magnet, an induced current due to the motion (§ 226) is set up in each molecule, and in such a direction that the molecules become elementary magnets, with their poles so directed towards the magnet in the field that there is repulsion between them. If this theory be true, it ought to be possible, as suggested by Maxwell, to lessen the intensity of magnetization of a body magnetized by induction, by increasing the strength of the field beyond a certain limit.

224. The Hall Effect.—Hitherto it has been assumed that when currents interact, it is their conductors alone which are affected, and that the currents in the conductors are not in any way altered. Hall has, however, discovered a fact which seems to show that currents may be displaced in their conductors. If the two poles of a voltaic battery be joined to two opposite arms of a cross of gold foil mounted on a glass plate, and if a galvanometer be joined to the other two arms at such points that no current flows through it, then if a magnet pole be brought opposite the face of the cross a permanent current will be indicated by the galvanometer. The same effect appears in the case of other metals. The direction of the permanent current and its amount differ under the same circumstances for different metals. The coefficient which represents the amount of the Hall effect in any metal is called the *rotational coefficient* of that metal.

Since the rotational coefficients of such metals as have been tested agree in sign and in relative magnitude with their thermo-electric powers (§ 235), it is argued by Bidwell, Ettinghausen, and others that the Hall effect is due to thermo-electric action.

225. Measurement of Current.—Instruments which are used to detect the presence of a current, or to measure its strength by means of the deflection of a magnetic needle, are commonly called *galvanometers.*

The simplest form of the galvanometer is the old instrument called the *Schweigger's multiplier.* It consists of a flat spool upon which an insulated wire is wound a number of times. The plane of the coils is vertical, and usually also coincides with the plane of the magnetic meridian. A magnetic needle is suspended in the interior of the spool. When a current is passed through the wire, the needle is deflected from the magnetic meridian. Usually, in order to make the indications of the apparatus more sensitive, a combination of two needles is used. They are joined rigidly together, so that when suspended the lower one hangs in the interior of the spool, and the other in the same plane directly above the spool. These needles are magnetized so that the positive end of one is above the negative end of the other. If they are of nearly equal strength, such a combination will have very little directive tendency in the earth's magnetic field. It is therefore called an *astatic* system. When a current passes in the wire, however, the lines of force due to the current form closed curves passing through the coil, and both needles tend to turn in the same direction. Since the earth's field offers almost no resistance to this tendency, an astatic system will indicate the presence of very feeble currents. The apparatus here described is no longer used to measure currents, but only to detect their presence and direction.

The *sine galvanometer* consists of a circular coil of insulated wire, set in the vertical plane, in the centre of which is a support for a magnetic needle. The needle can turn in the horizontal plane. When a current is sent through the coil, the magnet is deflected. The coil is then turned about the vertical axis, until the magnet lies in the plane of the coils. When this is the case, the equilibrium of the needle is due to the equality of the couples set up by the current in the coils and by the horizontal component of the earth's magnetism. The couple due to the horizontal component (Fig. 68) is $Hml \sin \phi$, where H represents the horizontal component, ml the magnetic moment of the magnet, and ϕ the angle made by the plane of the coils with the magnetic meridian. The couple due to the current is, by Biot's law, proportional to the current. It may then be set equal to $kmil$, where k is a constant factor depending upon the dimensions of the galvanometer. Since these two couples are equal, we have the equation

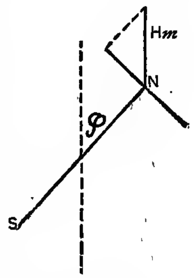

FIG. 68.

$$i = \frac{H}{k} \sin \phi. \tag{95}$$

With the same galvanometer, then, different currents are proportional to the sines of the angles made with the magnetic meridian by the plane of the coils when the needle lies in that plane. If i be greater than $\frac{H}{k}$, the equilibrium supposed in this explanation cannot occur.

The *tangent galvanometer* is that form of galvanometer which is commonly used to measure current in electromagnetic units. It can best be discussed by considering first the action of a single circular current of strength i upon a magnet pole situated at any point on the normal to the plane of the circle drawn from its centre.

The force due to any current element s (Fig. 69) upon the magnet pole m, at the distance l, is, by Biot's law, $\frac{mis}{l^2}$. This force tends to move the pole m at right angles to the plane containing s and the line joining s and m. If we represent by θ the angle between the line joining s and m and the x axis, the components of this force become $\frac{mis}{l^2}\sin\theta$ along the x axis, and $\frac{mis}{l^2}\cos\theta$ normal to the x axis. The equal element s', diametrically opposite s, also gives rise to two components, $\frac{mis'}{l^2}\sin\theta$ along the x axis, which is added to the similar component due to s, and $\frac{mis'}{l^2}\cos\theta$ normal to the x axis, which is opposed to and annuls the similar component due to s. Every other similar pair of elements will give rise to two similar components along the x axis, and will annul one another's action normal to the x axis. The total force on m will then be a force along the x axis equal to the sum of all the components along that axis.

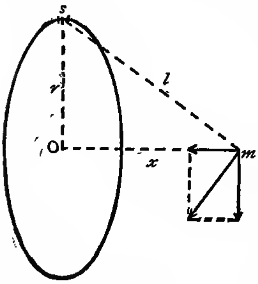

FIG. 69.

or $\Sigma\frac{mis}{l^2}\sin\theta$. This equals $\frac{2\pi mir}{l^2}\sin\theta$, where r is the radius of the circle. Since $\frac{r}{l} = \sin\theta$, this force may be written

$$\frac{2\pi mir^2}{l^3} = \frac{2\pi mir^2}{(x^2 + r^2)^{\frac{3}{2}}}.$$

If the circular coil considered be set vertical in the plane of the magnetic meridian, and a short magnetic needle be mounted at the point m, so as to turn in the horizontal plane, the needle will be deflected from the meridian, and will rest in equilibrium between the force due to the current and that due to the earth's magnetism. If the needle be so short that the distance of its poles from the x axis may be neglected, the formula just obtained will give the force upon its poles. Let l represent the half length of the needle (Fig. 70), ϕ its angle of deviation from the magnetic meridian, and d the distance from its centre to the plane of the coil. Then $d - l\sin\phi$ and $d + l\sin\phi$ represent the distances of the magnet's poles from the plane of the coil. The forces acting on these poles are then

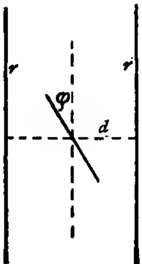

FIG. 70.

$$\frac{2\pi mir^2}{(r^2 + (d - l\sin\phi)^2)^{\frac{3}{2}}} \text{ and } \frac{2\pi mir^2}{(r^2 + (d + l\sin\phi)^2)^{\frac{3}{2}}}.$$

If another precisely similar coil be set at the same distance d from the point of suspension of the needle, on the opposite side of it, and if the current be sent through it in the same direction, two other forces equal to those just stated will act upon the needle, tending to turn it in the same direction. There will thus arise two couples with moments equal to

$$\frac{4\pi mir^2 l\cos\phi}{(r^2 + (d - l\sin\phi)^2)^{\frac{3}{2}}} \text{ and } \frac{4\pi mir^2 l\cos\phi}{(r^2 + (d + l\sin\phi)^2)^{\frac{3}{2}}},$$

both tending to turn the magnet in the same direction. The factors $\frac{1}{(r^2 + (d \pm l \sin \phi)^2)^{\frac{3}{2}}}$ are equal to

$$(r^2 + d^2)^{-\frac{3}{2}} \mp \tfrac{3}{2}(r^2 + d^2)^{-\frac{5}{2}}(2dl \sin \phi \pm l^2 \sin^2 \phi) \\ + \tfrac{15}{2}(r^2 + d^2)^{-\frac{7}{2}} d^2 l^2 \sin^2 \phi,$$

if we neglect all terms containing higher powers of l than the second. In this expression the upper or the lower signs must be used throughout. When we add the two moments of couple, we obtain for the total moment of couple acting on the needle the expression, after reduction,

$$\frac{8\pi mir^2 l \cos \phi}{(r^2 + d^2)^{\frac{3}{2}}}\left(1 - \frac{3}{2}\frac{(r^2 - 4d^2)}{(r^2 + d^2)^2} l^2 \sin^2 \phi\right).$$

This moment of couple is equal to that due to the horizontal intensity of the earth's magnetism, or $2mHl \sin \phi$. Setting these expressions equal, we obtain for i, if we neglect powers of l higher than the second,

$$i = \frac{H \tan \phi}{4\pi r^2}(r^2 + d^2)^{\frac{3}{2}}\left(1 + \frac{3}{2}\frac{(r^2 - 4d^2)}{(r^2 + d^2)^2} l^2 \sin^2 \phi\right). \quad (96)$$

The best form of the tangent galvanometer is so constructed that $d = \frac{r}{2}$. In this case the second term in the parenthesis disappears, and we have $i = \frac{5^{\frac{3}{2}}}{32} \cdot \frac{Hr}{\pi} \tan \phi$. The current is proportional to the tangent of the angle of deflection. If the galvanometer coils contain a number of turns equal in each coil to $\frac{n}{2}$, the proportion of the breadth to the depth of the coils may be so determined that the current is given by the equation

$$i = \frac{5^{\frac{3}{2}}}{16} \cdot \frac{HR}{\pi n} \tan \phi. \qquad (97)$$

In this equation R is the mean radius of the coil. All the quantities in this expression for i, except H, are either numbers or lengths, and H can be measured in absolute units. The tangent galvanometer can therefore be used to measure current in absolute units.

Weber's *electro-dynamometer* is an instrument with fixed coils like those of the tangent galvanometer, but with a small suspended coil substituted for the magnet. The small coil is usually suspended by the two fine wires through which the current is introduced into it, and the moment of torsion of this so-called *bifilar suspension* enters into the expression for the current strength. The same current is sent through the fixed and the movable coils, and a measurement of its strength can be obtained in absolute units, as with the tangent galvanometer. By a proper series of experiments, this measurement is made independent of the horizontal intensity of the earth's magnetism. When the current is reversed in the instrument, the couple tending to turn the suspended coil does not change. If the effects of terrestrial magnetism can be avoided, the electro-dynamometer can therefore be used to measure rapidly alternating currents.

226. Induced Currents.—It was shown in § 206 that the movement of a magnet in the neighborhood of a closed circuit will give rise, in general, to an electromotive force in the circuit, and that the current due to this electromotive force will be in the direction opposite to that current which, by its action upon the magnet, would assist the actual motion of the magnet. This current is called an *induced current*. From the equivalence between a magnetic shell and an electrical current, it is plain that a similar induced current will be produced in a closed circuit by the movement near it of an electrical current or any part of one. Since the joining up or breaking the circuit carrying a current is equivalent to bringing up that same current from an infinite distance, or removing it to an infinite distance, it is further evident that similar induced currents will be produced in a closed circuit when a circuit is made or broken in its presence.

The demonstration of the production of induced currents in § 206 depends upon the assumption that the path of the magnet pole is such that work is done upon it by the current assumed to exist in the circuit. The potential of the magnet pole relative to the current is changed.

The change in potential from one point to another in the magnetic field due to a closed current is (Eq. 93) equal to $i(\omega_{,} - \omega + 4\pi n)$, and the work done on a magnet pole m, in moving it from one point to another, is $mi(\omega_{,} - \omega + 4\pi n)$. In the demonstration of § 206 we may substitute $m(\omega_{,} - \omega + 4\pi n)$ for A, and, provided the change in potential be uniform, we obtain at once the expression $-\dfrac{m(\omega_1 - \omega + 4\pi n)}{t}$ for the electromotive force due to the movement of the magnet pole. If the change in potential be not uniform, we may conceive the time in which it occurs to be divided into indefinitely small intervals, during any one of which, t, it may be considered uniform. Then the limit of the expression $-\dfrac{m(\omega_1 - \omega + 4\pi n)}{t}$, as t becomes indefinitely small, is the electromotive force during that interval.

The current strength due to this electromotive force is

$$i_{,} = -\frac{m(\omega_{,} - \omega + 4\pi n)}{Rt}.$$

If the induced current be steady, the total quantity of electricity flowing in the circuit is expressed by

$$i_{,}t = -\frac{m(\omega_{,} - \omega + 4\pi n)}{R}.$$

The total quantity of electricity flowing in the circuit depends, therefore, only upon the initial and final positions of the magnet pole, and the number of times it passes through the circuit, and not upon its rate of motion. The electromotive force due to the movement of the magnet, and consequently the current strength, depends, on the other hand, upon the rate at which the potential changes with respect to time.

A more general statement, which will include all cases of the production of induced currents, may be derived by the use of the method of discussion given in § 219. The change in potential of a closed circuit, carrying a current in a magnetic field, may be measured by the change in the number of lines of force which pass through it in the positive direction. Any movement which changes the number of lines of force will set up in the circuit an electromotive force, and an induced current in a sense opposite to that current which would by its action assist the movement. As in the elementary case which has just been discussed, the total quantity of electricity passing in the circuit depends only upon the total change in the number of lines of force passing through the circuit in the positive direction, but the electromotive force and current strength depend on the rate of change in the number of lines of force.

It is often convenient, especially when considering the movement of part of a circuit in a magnetic field, to speak of the change in the number of lines of force enclosed by the circuit as the number of lines of force cut by the moving part of the circuit. The direction of the induced current in the moving part of the circuit, if it be supposed to move normal to the lines of force, is related to the direction of motion and to the positive direction of the lines of force cut, in such a way that the three directions may be represented by the positive directions of the three co-ordinate axes of x, y, and z, when the x axis represents the direction of motion, the y axis the lines of magnetic force, and the z axis the direction of the induced current. The positive directions of the three axes is such that, if we rotate the positive x axis through a right angle about the z axis, clockwise as seen by one looking along the positive direction of the z axis, it will coincide with the positive y axis.

The fact that induced currents are produced in a closed circuit by a variation in the number of lines of magnetic force included in it was first shown experimentally by Faraday in 1831. He placed one wire coil, in circuit with a voltaic battery, inside another which was joined with a sensitive galvanometer. The first he called the primary, the second the secondary, circuit. When the battery circuit was made or broken, deflections of the galvanometer were observed. These were in such a direction as to indicate a current in the secondary coil, when the primary circuit was made, in the opposite direction to that in the primary, and when the primary circuit was broken, in the same direction as that in the primary. When the positive pole of a bar magnet was thrust into or withdrawn from the secondary coil, the galvanometer was deflected. The currents indicated were related to the direction of motion of the positive magnet pole, as the directions of rotation and propulsion in a left-handed screw. The direction of the induced currents in these experiments is easily seen to be in accordance with the law above stated, that the induced currents are always in the opposite direction to those currents which would, by their action, assist the motion.

This law of induced currents in its general form was first announced by Lenz in 1834, soon after Faraday's discovery of the production of induced currents. It is known as *Lenz's law*.

The case in which an induced current in the secondary circuit is set up by making the primary circuit is, as has been said, an extreme case of the movement of the primary circuit from an infinite distance into the presence of the secondary. The experiments of Faraday and others show that the total quantity of electricity induced when the primary circuit is made is exactly equal and opposite to that induced when the primary circuit is broken. They also show that the electromotive force induced in the secondary circuit is independent of the materials constituting either circuit, and is proportional to the current strength in the primary circuit. These results are consistent with the formula already deduced for the induced current.

227. Self-induction.—When a current is set up in any circuit, the different parts of the circuit act on one another in the relation of primary and secondary circuits. In a long straight wire, for example, the current which is set up through any small area in the cross-section of the wire tends to develop an opposing electromotive force through every other area in the same cross-section. The true current will thus be temporarily weakened, and will require a certain time to attain its full strength. On the other hand, when the circuit is broken, the induced electromotive force is in the same direction as the electromotive force of the circuit. Since the time occupied by the change of the true current from its full value to zero, when the circuit is broken, is very small, the induced electromotive force is very great. The current formed at breaking is called the *extra current*, and gives rise to a spark at the point where the circuit is broken. The extra current may be heightened by anything which will increase the change in the number of lines of force, as by winding the wire in a coil and by inserting in the coil a piece of soft iron. This action of a circuit on itself is called *self-induction*.

228. Electromagnetic Unit of Electromotive Force.—If the circuit considered in § 226 move from a point where its potential relative to the magnet pole is $m\omega_{,}$ to one where it is $m\omega$, provided that the magnetic pole do not pass through the circuit, and that the movement be so carried out that the induced current is constant, the electromotive force of the induced current is $-\frac{m(\omega_{,} - \omega)}{t}$. If the movement take place in unit time, and if $m(\omega_{,} - \omega)$ also equal unity, the electromotive force in the circuit is defined to be *unit electromotive force.*

The expression $m(\omega_{,} - \omega)$ is equivalent to the change in the number of lines of force passing through the circuit in the positive direction. More generally, then, if a circuit or part of a circuit so move in a magnetic field that, in unit time, the number of lines of force passing through the circuit in the positive direction increase or diminish by unity, at a uniform rate, the electromotive force induced is unit electromotive force.

The simplest way in which these conditions can be presented is as follows: Suppose two parallel straight conductors at unit distance apart, joined at one end by a fixed cross-piece. Suppose the circuit to be completed by a straight cross-piece of unit length which can slide freely on the two long conductors. Suppose this system placed in a magnetic field of unit intensity, so that the lines of force are everywhere perpendicular to the plane of the conductors. Then, if we suppose the sliding piece to be moved with unit velocity perpendicular to itself along the parallel conductors, the electromotive force set up in the circuit will be the unit electromotive force.

The unit of electromotive force thus defined is the electromagnetic unit. In practice another unit is used, called the *volt.* It contains 10^8 C. G. S. electromagnetic units.

To obtain the dimensions of electromotive force in the electromagnetic system we need first the dimensions of number of lines of force. From the convention adopted by which lines of force are used to measure the strength of a magnetic field we have $\left[\frac{n}{L^2}\right] = [H]$; whence $[n] = M^{\frac{1}{2}} L^{\frac{3}{2}} T^{-1}$. Since the electromotive force is measured by the rate of change of the number of lines of force we have $[e] = \left[\frac{n}{T}\right] = M^{\frac{1}{2}} L^{\frac{3}{2}} T^{-2}$.

The definition of electromotive force is consistent, as it must be, with the equation $ie =$ rate of work, or work divided by time. This equation is the same as that discussed in § 202, and holds whichever system of units is adopted. In the determination of the unit of electromotive force the arrangement given above is, of course, impracticable. In those experiments which have been made, the induced electromotive force which was due to the rotation of a circular coil in a magnetic field was determined by calculation.

229. Apparatus employing Induced Currents.—The production of induced currents by the relative movements of conductors and magnets is taken advantage of in the construction of pieces of apparatus which are of great importance not only for laboratory use but in the arts.

The *telephonic receiver* consists essentially of a bar magnet around one end of which is carried a coil of fine insulated wire. In front of this coil is placed a thin plate of soft iron. When the coils of two such instruments are joined in circuit by conducting wires, any disturbance of the iron diaphragm in front of one coil will change the magnetic field near it, and a current will be set up in the circuit. The strength of the magnet in the other instrument will be altered by this current, and the diaphragm in front of it will move. When the diaphragm of the first instrument, or *transmitter*, is set in motion by sound-waves due to the voice, the induced currents, and the consequent movements of the diaphragm of the second instrument, or *receiver*, are such that the words spoken into the one can be recognized by a listener at the other.

Other transmitters are generally used, in which the diaphragm presses upon a small button of carbon. A current is passed from a battery through the diaphragm, the carbon button, and the rest of the circuit, including the receiver. When the diaphragm moves, it presses upon the carbon button and alters the resistance of the circuit at the point of contact. This change in resistance gives rise to a change in the current, and the diaphragm of the receiver is moved. The telephone serves in the laboratory as a most delicate means of detecting a change of current in a circuit.

The various forms of magneto-electrical and dynamo-electrical machines are too numerous and too complicated for description. In all of them an arrangement of conductors, usually called the *armature*, is moved in a powerful magnetic field, and a suitable arrangement is made by which the currents thus induced may be led off and utilized in an outside circuit. The magnetic field is sometimes established by permanent magnets, and the machine is called a *magneto-machine*. In most cases, however, the circuit containing the armature also contains the coils of the electromagnets to which the magnetic field is due. When the armature rotates, a current starts in it, at first due to the residual magnetism of some part of the machine: this current passes through the field magnets and increases the strength of the magnetic field. This in turn reacts upon the armature, and the current rapidly increases until it attains a maximum due to the fact that the magnetic field does not increase proportionally to the current which produces it. Such a machine is called a *dynamo-machine*.

The *induction coil*, or *Ruhmkorff's coil*, consists of two circuits wound on two concentric cylindrical spools. The inner or primary circuit is made up of a comparatively few layers of large wire, and the outer, or secondary, of a great number of turns of fine wire. Within the primary circuit is a bundle of iron wires, which, by its magnetic action, increases the electromotive force of the induced current in the secondary coil. Some device is employed by which the primary circuit can be made or broken mechanically. The electromotive force of the induced current is proportional to the number of windings in the secondary coil, and as this is very great the electromotive force of the induced current greatly exceeds that of the primary current. The electromotive force of the induced current set up when the primary circuit is broken is further heightened by a device proposed by Fizeau. To two points in the primary circuit, one on either side of the point where the circuit is broken, are joined the two surfaces of a condenser. When the circuit is broken, the extra current, if the condenser be not introduced, forms a long spark across the gap and so prolongs the fall of the primary current to zero. The electromotive force of the induced current is therefore not so great as it would be if the fall of the primary current could be made more rapid. When the condenser is introduced, the extra current is partly spent in charging the condenser, the difference of potential between the two sides of the gap is not so great, the length of the spark and consequently the time taken by the primary current to become zero is lessened, and the electromotive force of the induced current is proportionally increased.

230. Resistance.—As in the discussion of § 203, we may here define the ratio of the electromotive force to the current in any circuit as the resistance in that circuit. The *electromagnetic unit of resistance* is the resistance of that circuit in which unit electromotive force gives rise to unit current, when both these quantities are measured in electromagnetic units. In the example given in § 228, if we insert a galvanometer in that part of the circuit occupied by the fixed cross-piece, and assume that the resistance of every part of the circuit except the sliding piece is zero, the resistance of the sliding piece will be unity when, moving with unit velocity, it gives rise to unit current in the galvanometer. If it move with any other velocity v, and still produce unit current in the galvanometer, its resistance will be numerically equal to the velocity v. For the electromotive force produced by a movement with that velocity is v, and the ratio of that electromotive force to unit current is v, which is the resistance by definition.

A unit of resistance, intended to be the C.G.S. electromagnetic unit, was determined by a committee of the British Association by the following method: A circular coil of wire, in the centre of which was suspended a small magnetic needle, was mounted so as to rotate with constant velocity about a vertical diameter. From the dimensions and velocity of rotation of the coil and the intensity of the earth's magnetic field, the induced electromotive force in the coil was calculated. The current in the same coil was determined by the deflection of the small magnet. The ratio of these two quantities gave the resistance of the coil.

In practice another unit of resistance is used, called the *ohm*. It would be the resistance of a sliding piece in the arrangement before described which would give rise to the C. G. S. unit of current if it were to move with a velocity of one billion centimetres in a second. The *true ohm* thus contains 10 C. G. S. electromagnetic units. The dimensions of resistance in the electromagnetic system are $[r] = \left[\frac{e}{i}\right] = LT^{-1}$. The dimensions of resistance are therefore those of a velocity, as might be inferred from the measure of resistance in terms of velocity in the example given above.

The standard of resistance, usually called the B. A. unit, determined by the committee of the British Association, has a resistance somewhat less than the true ohm as it is here defined. In practical work resistances are used which have been compared with this standard. The Electrical Congress of 1884 defined the *legal ohm* to be "the resistance of a column of mercury of one square millimetre section and of 106 centimetres of length at the temperance of freezing." The legal ohm contains 1.0112 B. A. units. Boxes containing coils of wire of definite resistance, so arranged that by different combinations of them any desired resistance may be introduced into a circuit, are called *resistance boxes* or *rheostats*.

231. Kirchhoff's Laws.—In circuits which are made up of several parts, forming what may be called a network of conductors, there exist relations between the electromotive forces, currents, and resistances in the different branches, which have been stated by Kirchhoff in a way which admits of easy application.

Several conventions are made with regard to the positive and negative directions of currents. In considering the currents meeting at any point, those currents are taken as positive which come up to the point, and those as negative which move away from it. In travelling around any closed portion of the network, those currents are taken as positive which are in the direction of motion, and those as negative which are opposite to the direction of motion. Further, those electromotive forces are positive which tend to set up a positive current in their respective branches. With those conventions *Kirchhoff's laws* may be stated as follows:

1. The algebraic sum of all the currents meeting at any point of junction of two or more branches is equal to zero. This first law is evident, because, after the current has become steady, there is no accumulation of electricity at the junctions.

2. The sum, taken around any number of branches forming a closed circuit, of the products of the currents in those branches into their respective resistances is equal to the sum of the electromotive forces in those branches. This law can easily be seen to be only a modified statement of Ohm's law, which was given in § 203.

These laws may be best illustrated by their application in a form of apparatus known as *Wheatstone's bridge*. The circuit of the Wheatstone's bridge is made up of six branches. An end of any branch meets two, and only two, ends of other branches, as shown in Fig. 71. In the branch 6 is a voltaic cell with an electromotive force E. In the branch 5 is a galvanometer which will indicate the presence of a current in that branch. In the other branches are conductors, the resistances of which may be called respectively r_1, r_2, r_3, r_4.

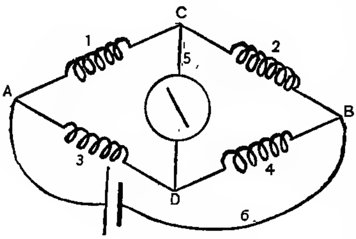

FIG. 71.

From Kirchhoff's first law the sum of the currents meeting at the point C is $i_1 + i_2 + i_5 = 0$, and of those meeting at the point D is $i_3 + i_4 + i_5 = 0$. By the second law, the sum of the products ir in the circuit ADC is $i_1r_1 + i_3r_3 + i_5r_5 = 0$, and in the circuit DBC is $i_4r_4 + i_2r_2 + i_5r_5 = 0$, since there are no electromotive forces in those circuits. If we so arrange the resistances of the branches 1, 2, 3, 4 that the galvanometer shows no deflection, then the current i_5 is zero, and these equations give the relations, $i_1 = -i_2$, $i_3 = -i_4$, $i_1r_1 = -i_3r_3$, $i_2r_2 = -i_4r_4$. From these four equations follows at once a relation between the resistances, expressed in the equation

$$r_1r_4 = r_2r_3. \tag{98}$$

If, therefore, we know the value of r_3 and know the ratio of r_1 to r_2, we may obtain the value of r_4.

This method of comparing resistances by means of the Wheatstone's bridge is of great importance in practice. By the use of a form of apparatus known as the *British Association bridge* the method can be carried to a high degree of accuracy. In this form of the bridge, the portion marked ACB (Fig. 71) is a straight cylindrical wire, along which the end of the branch CD

is moved until a point C is found, such that the galvanometer shows no deflection. The two portions of the wire between C and A, and C and B, are then the two conductors of which the resistances are r_1 and r_2, and these resistances are proportional to the lengths of those portions (§ 204). The ratio of r_1 to r_2 is therefore the ratio of the lengths of wire on either side of C, and only the resistance of r_3 need be known in order to obtain that of r_4.

It is often convenient in determining the relations of current and resistance in a network of conductors to use Ohm's law (§ 203), directly, and consider the difference of potential between the two points on a conductor as equal to the product ir. When a part of a circuit is made up of several portions which all meet at two points A and B, the relation between the whole resistance and that of the separate parts may be obtained easily in this way. For convenience in illustration we will suppose the divided circuit (Fig. 72) made up of only three portions, 1, 2, 3, meeting at the points A and B, and that no electromotive force exists in those portions. Then the difference of potential between A and B is $V_A - V_B = i_1 r_1 = i_2 r_2 = i_3 r_3$. We have also by Kirchhoff's first law $-i_4 = i_1 + i_2 + i_3$. By the combination of these equations we obtain

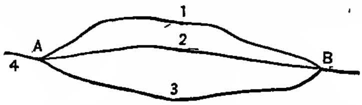

FIG. 72.

$$-i_4 = (V_A - V_B)\left(\frac{1}{r_1} + \frac{1}{r_2} + \frac{1}{r_3}\right). \tag{99}$$

The current in the divided circuit equals the difference of potential between A and B multiplied by the sum of the reciprocals of the resistances of the separate portions. If we set this sum equal to $\frac{1}{r}$, and call r the resistance of the divided circuit, we may say that the reciprocal of the resistance of a divided circuit is equal to the sum of the reciprocals of the resistances of the separate portions of the circuit. When there are only two portions into which the circuit is divided, one of them is usually called a *shunt*, and the circuit a shunt circuit.

An arrangement devised by Clark, called the *Clark's potentiometer*, used to compare the electromotive forces of voltaic cells, depends for its action on the principles here discussed. It consists of a spiral of evenly drawn wire coiled about a rubber cylinder, with arrangements by which contact can be made with it at both ends and at any point along it. Let us call the cells to be compared cell 1 and cell 2, and let the electromotive force of cell 1 be the greater. To the two ends of the spiral are joined the terminals of a circuit which we will call A, containing a constant voltaic battery, of which the electromotive force is greater than that of either cell 1 or cell 2, and a set of resistances which can be varied. To the same points are joined the terminals of a circuit which we will call B, containing cell 1, and a sensitive galvanometer. The positive poles of the constant battery and of cell 1 are joined to the same end of the spiral. The resistance is then modified in circuit A until the galvanometer in circuit B shows no deflection. The difference of potential between the ends of the spiral is, therefore, equal and in the opposite direction to the electromotive force of cell 1. The positive pole of cell 2 is now joined to the end of the spiral to which the positive poles of the other circuits are joined, and with the free end of a circuit C, containing cell 2 and a sensitive galvanometer, contact is made at different points on the spiral until the point is found at which, when contact is made, the galvanometer in C shows no deflection. The difference of potential between that point and the end of the spiral joined to the positive poles is equal and opposite to the electromotive force of cell 2. The electromotive forces of the two cells are then proportional to the lengths of the wire between the points of contact of their terminals; that is, the electromotive force of cell 1 is to that of cell 2 as the length of the wire spiral is to that portion of its length between the two terminals of cell 2. For, since the wire is uniform, its resistance is proportional to its length, and if we represent the potential of the common point of contact of the positive poles by V, the potentials of the points of contact of the two negative poles by V_1 and V_2, the current in the spiral by i, and the resistances of the lengths of wire considered by r_1 and r_2, we have

$$i = \frac{V_1 - V}{r_1} = \frac{V_2 - V}{r_2}.$$

The rules for joining up sets of voltaic cells in circuits so as to accomplish any desired purpose may be discussed by the same method. Let us suppose that there are n cells, each with an electromotive force e and an internal resistance r, and that the external resistance of the circuit is s. If m be a factor of n, and if we join up the cells with the external resistance so as to form a divided circuit of m parallel branches, each containing $\frac{n}{m}$ cells, we shall have for the electromotive force in such a circuit $\frac{ne}{m}$, and for the resistance of the circuit $s + \frac{nr}{m^2}$. The current in the circuit is therefore $i = \frac{mne}{m^2s + nr}$. Two cases may arise which are common in practice. The resistance s of the external circuit may be so great that, in comparison with m^2s, nr may be neglected. In that case i is a maximum when $m = 1$, that is, when the cells are arranged *tandem*, or in *series*, with their unlike poles connected. On the other hand, if m^2s be very small as compared with nr, it may be neglected, and i becomes a maximum when $m = n$, that is, when the cells are arranged *abreast*, or in *multiple arc*, with their like poles in contact.

232. Ratio between the Electrostatic and Electromagnetic Units.—When the dimensions of any electrical quantity derived from its electrostatic definition are compared with its dimensions derived from its electromagnetic definition, the ratio between them is always of the dimensions of some power of a velocity. The ratio between the electrostatic and electromagnetic unit of any electrical quantity is, therefore, of the dimensions of some power of a velocity. If, therefore, this ratio be obtained for any set of units, the number expressing it will also express some power of a velocity. This velocity is an absolute quantity or constant of nature. Whatever changes are made in the units of length and time, the number expressing this velocity in the new units will also express the ratio of the two sets of electrical units.

This ratio, which is called v, can be measured in several ways.

The first method, used by Weber and Kohlrausch, depends upon the comparison of a quantity of electricity measured in the two systems. From the dimensions of current in the electromagnetic system we have the dimensions of quantity $[q] = [iT] = M^{\frac{1}{2}}L^{\frac{1}{2}}$. The dimensions of quantity in the electrostatic system are $[Q] = M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-1}$. The ratio of these dimensions is $\left[\frac{Q}{q}\right] = LT^{-1}$, or, the number of electrostatic units of quantity in one electromagnetic unit is the velocity v.

In Weber and Kohlrausch's method the charge of a Leyden jar was measured in electrostatic units by a determination of its capacity and the difference of potential between its coatings. The current produced by its discharge through a galvanometer was used to measure the same quantity in electromagnetic measure.

Thomson determined v by a comparison of an electromotive force measured in the two systems. He sent a current through a coil of very high known resistance, and measured it by an electro-dynamometer. The electromagnetic difference of potential between the two ends of the resistance coil was then equal to the product of the current by the resistance. The electrostatic difference of potential between the same two points was measured by an absolute electrometer. From the dimensional formulas we have

$$\left[\frac{E}{e}\right] = \frac{M^{\frac{1}{2}}L^{\frac{1}{2}}T^{-1}}{M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-2}} = L^{-1}T.$$

The number of electromagnetic units of electromotive force in one electrostatic unit is v. The ratio of the numbers expressing the electromagnetic and the electrostatic measures of the electromotive force in Thomson's experiment is therefore the quantity v. This experiment was carried out by Maxwell in a different form, in which the electrostatic repulsion of two similarly charged disks was balanced by an electromagnetic attraction between currents passing through flat coils on the back of the two disks.

Other methods, depending on comparisons of currents, of resistances, and other electrical quantities, have been employed. The methods described are historically interesting as being the first ones used. The values of v obtained by them differed rather widely from one another. Recent determinations, however, give more consistent results. It is found that v, considered as a velocity, is about $3 \cdot 10^{10}$ centimetres in a second. This velocity agrees very closely with the velocity of light.

The physical significance of this quantity v may be understood from an experiment of Rowland. The principle of the experiment is as follows. If we consider an indefinitely extended plane surface on which the surface density of electrification is σ, measured in electrostatic units, or $\frac{\sigma}{v}$ measured in electromagnetic units, since the ratio of the electrostatic to the electromagnetic unit of quantity is v; and conceive it to move in its own plane with a velocity x; the charge moving with it may be considered as the equivalent of a current in that surface, the strength of which, measured by the quantity of electricity which crosses a line of unit length, perpendicular to the direction of movement, in unit time, is $\frac{\sigma x}{v}$. The force due to such a current on a magnet may be calculated. Conversely, if the force on the magnet be observed, and the surface density σ and the velocity x be also measured, the value of v may be calculated. The probability of such an action as the one here described was stated by Maxwell.

The experiment by which Rowland verified Maxwell's view consisted in rotating a disk cut into numerous sectors, each of which was electrified, under an astatic magnetic needle. During the rotation of the disk, a deflection of the needle was observed, in the same sense as that in which it would have moved if a current had been flowing about the disk in the direction of its rotation. From the measured values of the deflecting force, of the surface density of electrification on the disk, and the velocity of rotation, Rowland calculated a value of v which lies between those given by Weber and Maxwell.

It may be seen that, if the velocity x of the moving surface which we at first considered be equal to v, the equivalent current strength in the surface will be σ. If we imagine another such surface near the one already considered, the repulsion between them due to their opposite charges is $2\pi\sigma^2$ for every unit of surface (§ 198). It can be shown, by a method too extended to be given here, that the attraction between two currents in the same surfaces, of which the strengths in the surface are both σ, is also expressed by $2\pi\sigma^2$ for every unit of surface. Hence if the surfaces, so charged that the surface density of their electrification is σ, can move with a velocity in their own planes equal to v, the repulsion of the charges will exactly counterbalance the attraction of the currents due to their movement.

CHAPTER VI.

THERMO-ELECTRIC RELATIONS OF THE CURRENT.

233. **Thermo-electric Currents.**—The heating or cooling of a junction of two dissimilar metals by the passage of a current, referred to in § 200 as the Peltier effect, is the reverse of a phenomenon discovered in 1822–23 by Seebeck. He found that, when the junction of two dissimilar metals was heated, a current was sent through any circuit of which they formed a part. It has since been shown that the same phenomenon appears if the junction of two liquids, or of a liquid and a metal, be heated. This fact, as has been already shown in § 206, follows as a result of the Peltier phenomenon. If we designate by P the heat developed at the junction by the passage of unit current for unit time, we may substitute it for the expression $\frac{A}{t}$ in the general equation of § 206, and obtain $I = \frac{E - P}{R}$. The counter electromotive force set up at the heated junction is the coefficient P, and is the measure of the true electromotive force of contact (§ 214). The contact electromotive force of Volta does not agree in magnitude and not always in sign with this electromotive force. From this fact it is evident that the contact electromotive force of Volta is at least partially due to the air or other medium in which the bodies which are tested are placed.

If the electromotive force E and the current I be reversed in the circuit, the junction is cooled and we obtain $I = \frac{E + P}{R}$.

The electromotive force at the junction, therefore, tends to increase the electromotive force of the circuit. Since this is opposite to the electromotive force of the circuit in the case in which the junction is heated, the direction of the electromotive force at the junction is the same as that found in the other case. If, then, there be no electromotive force E in the circuit, we have $I = -\frac{P}{R}$ in case a unit of heat is communicated to the junction and absorbed by it in unit time, and $I = \frac{P}{R}$ in case a similar quantity of heat is removed from the junction by cooling.

If two strips of dissimilar metals, for example antimony and bismuth, be placed side by side, and united at one end of the pair, being everywhere else insulated from one another, the combination is called a *thermo-electric element.* If several such elements be joined in series, so that their alternate junctions lie near together and in one plane, as indicated in Fig. 73, such an arrangement is called a *thermo-pile.* When one face of the pile is heated, the electromotive force of the pile is the sum of the electromotive forces of the several elements. Such an instrument was used by Melloni, in connection with a delicate galvanometer, in his researches on radiant heat.

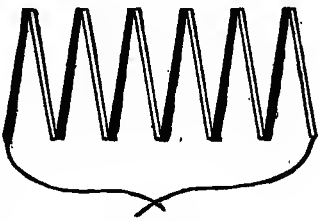

FIG. 73.

When a thermo-electric element is constructed of any two metals, that metal is said to be *thermo-electrically positive* to the other from which the current flows across the heated junction.

234. **Thermo-electric Series.**—It was found by the experiments of Seebeck himself, and those of others, that the metals may be arranged in a series such that any metal in it is thermo-

electrically positive to those which follow it, and thermo-electrically negative to those which precede it.

If a circuit be formed of any two metals in this series, and one of the junctions be kept at the temperature zero, while the other is heated to a fixed temperature, there will be set up an electromotive force which can be measured. If now the circuit be broken at either junction, and the gap filled by the introduction of any other metals of the series, then, provided that the junction which has not been disturbed be kept at the temperature which it previously had, and that the other junctions in the circuit be all raised to the temperature of the junction which was broken, there will be the same electromotive force in the circuit as existed before the introduction of the other metals of the series. It is manifest, then, that in a circuit made up of any metals whatever, at one temperature, no electromotive force can be set up by changing the temperature of the circuit as a whole.

Thomson showed that it is not necessary for the production of thermal currents that the circuit should contain two metals; but that want of homogeneity arising from any strain of one part of an otherwise homogeneous circuit will also admit of the production of such currents. It has also been shown that when a portion of an iron wire is magnetized, and is heated near one of the poles produced, a thermal current will be set up.

Cumming discovered in 1823 that, if the temperature of one junction of a circuit of two metals be gradually raised, the current produced will increase to a maximum, then decrease until it becomes zero, after which it is reversed and flows in the opposite direction. The experiments of Avenarius, Tait, and Le Roux show that, for almost all metals, the temperature of the hot junction at which the maximum current occurs is the mean between the temperatures of the two junctions at which the current is reversed.

235. Thermo-electric Diagram.—The facts hitherto discovered in relation to thermo-electricity may be collected in a general formula or exhibited by means of a thermo-electric diagram.

Let us consider a circuit of two metals, copper and lead, in which both junctions are at first at the same temperature. We may assume that there is an equal electromotive force of contact at both junctions acting from lead to copper. If one of the junctions be gradually heated, a current will be set up, passing from lead to copper across the hot junction. The heating has disturbed the equilibrium of electromotive forces, and has increased the electromotive force across the hot junction from lead to copper. The rate at which this electromotive force changes with change in the temperature is called the ***thermo-electric power*** of the two metals. That is, if E represent the electromotive force, t the temperature, and θ the thermo-electric power, we have $\frac{E_1 - E_0}{t_1 - t_0} = \theta_1$, in the limit where t_1 and t_0 are indefinitely near one another. Hence if we lay off on the axis of abscissas (Fig. 74) an infinitesimal length $t_1 - t_0$, and erect as ordinate the corresponding thermo-electric power θ_1, the area of the rectangle formed by the two lines will represent the electromotive force $E_1 - E_0$, due to the change in temperature. If, beginning at the point t_1, we lay off the similar infinitesimal length $t_2 - t_1$, and erect as ordinate the thermo-electric power θ_2, we shall obtain another rectangle representing the electromotive force $E_2 - E_1$. So for any temperature changes the total area of the figure bounded by the axis of temperatures, by the ordinates representing the thermo-electric powers at the temperatures t_0 and t_x, and by the curve AA' passing through the summits of the rectangles so obtained, will represent the electromotive force due to the heating of the junction from t_0 to t_x.

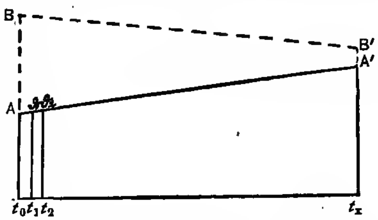

FIG. 74.

It was found by Tait and Le Roux that the thermo-electric power, referred to lead as a standard, of all metals but iron and nickel, is proportional to the rise in temperature. The curve AA' is therefore for those metals a straight line. For iron and nickel the curve is not straight.

For another metal in comparison with lead, the line BB', corresponding to the line AA' for copper, may have a different direction. From what has been said about the possibility of arranging the metals in a thermo-electric series, it is evident that the thermo-electric power between copper and the other metal is the difference of their thermo-electric powers referred to lead, and that the electromotive force at the junction of the two metals, due to a rise of temperature from t_0 to t_x, is represented by the area of the figure contained by the two terminal ordinates and the two lines AA' and BB'. The thermo-electric power is reckoned positive when the current sets from lead to copper across the hot junction. In the diagram the thermo-electric power $A'B'$ is positive, and the electromotive force indicated by the area is from copper to the other metal across the hot junction. At the point where the lines AA' and BB' intersect, the thermo-electric power for the two metals vanishes. The temperature at which this occurs is called the ***neutral temperature*** and is designated by t_n. When the temperature t_x lies on the other side of the neutral temperature from t_0, the thermo-electric power becomes negative, and the electromotive force due to the rise in temperaature from t_n to t_x is negative. In Fig. 75 it is at once seen that $A'B'$ is negative for t_x, and that the area $NA'B'$ is also negative. The electromotive force due to a rise of temperature from t_0 increases until the temperature of the hot junction is t_n, when it is a maximum, and then de-

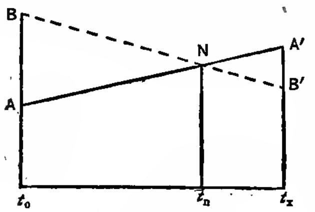

FIG. 75.

creases. When the area $NA'B'$ becomes equal to the area ANB, the total electromotive force is zero; when $NA'B'$ is greater than ANB, the electromotive force becomes negative, and the current is reversed. In case AA' and BB' are straight lines it is plain that the temperature t_x, at which this reversal occurs, will be such that the neutral temperature t_n is a mean between t_0 and t_x.

The same facts can be represented by a general formula. Thomson first pointed out that the fact of thermo-electrical inversion necessitates the view that the thermo-electric power at a junction is a function of the temperature of that junction. Avenarius embodied this idea in a formula, which his own researches, and those of Tait, show to be closely in agreement with experiment. Let us call the hot junction 1 and the cool junction 2, and set the electromotive force at each junction as a quadratic function of the absolute temperatures. We have $E_1 = A + bt_1 + ct_1^2$ and $E_2 = A + bt_2 + ct_2^2$, where A, b, and c are constants. The difference $E_1 - E_2$, or the electromotive force in the circuit, is

$$\begin{aligned} E_1 - E_2 &= b(t_1 - t_2) + c(t_1^2 - t_2^2) \\ &= (t_1 - t_2)(b + c(t_1 + t_2)) \end{aligned}$$

This equation may be put in the form used by Tait, if we write $b = at_n$ and $c = -\dfrac{a}{2}$. We then have

$$E_1 - E_2 = a(t_1 - t_2)(t_n - \tfrac{1}{2}(t_1 + t_2)) \tag{100}$$

The electromotive force in the circuit can become zero when either of these terms equals zero. It has been already stated that when $t_1 = t_2$, or when both junctions are at the same temperature, there is no electromotive force in the circuit.

When $\frac{1}{2}(t_1 + t_2)$ equals t_n, or when the mean of the temperatures of the hot and cold junctions equals a certain temperature, constant for each pair of metals, there will be also no electromotive force in the circuit. This temperature t_n is that which has already been called the neutral temperature. The formula also assigns the value to that temperature t_1 at which, for fixed values of t_n and t_2, the electromotive force in the circuit is a maximum. If we represent the difference between t_n and t_1 by x, then $t_1 = t_n \pm x$. Using this value in the formula, we obtain $E_1 - E_2 = \frac{a}{2}((t_n - t_2)^2 - x^2)$. This is manifestly a maximum when $x = 0$. The electromotive force in a circuit is then, according to the formula, a maximum when the temperature of one junction is the neutral temperature.

The formula also shows that the thermo-electric power is zero when $t_1 = t_n$. We may set $E_1 = A + at_nt_1 - \frac{a}{2}t_1^2$. Now if t_1 take any small increment Δt_1, E_1 has a corresponding increment ΔE_1. Hence we have

$$E_1 + \Delta E_1 = A + at_nt_1 - \frac{a}{2}t_1^2 + at_n\Delta t_1 - at_1\Delta t_1,$$

if we neglect the term containing Δt_1^2. From this equation we obtain $\frac{\Delta E_1}{\Delta t_1} = at_n - at_1$, which in the limit, as Δt_1 becomes indefinitely small, is the thermo-electric power at the temperature t_1. It is positive for values of t_1 below t_n; is zero for $t_1 = t_n$, and negative for higher values of t_1. That is, if we assume $t_1 = t_2$ lower than t_n, and then gradually raise the temperature t_1, the thermo-electric power at the heated junction is at first positive, but continually decreases in numerical value, until at $t_1 = t_n$ it becomes zero. At that temperature, then, the metals are thermo-electrically neutral to one another, and a small change in the temperature does not change the electromotive force at the junction.

236. The Thomson Effect.—Thomson has shown that, in certain metals, there must be a reversible thermal effect when the current passes between two unequally heated parts of the same metal. Let us suppose a circuit of copper and iron, of which one junction is at the neutral temperature, and the other below the neutral temperature. The current then sets from copper to iron across the hot junction. In the hot junction there is no thermal effect produced, because the metals are at the neutral temperature. Across the cold junction the current is flowing from iron to copper, and hence is evolving heat. The current in the circuit can be made to do work, and since no other energy is imparted to the circuit this work must be done at the expense of the heat in the circuit. Since heat is not absorbed at either junction, it must be absorbed in the unequally heated parts of the circuit between the junctions.

To show this, Thomson used a conductor the ends of which were kept at constant temperatures in two coolers, while the central portion was heated. When a current was passed through this conductor, thermometers, placed in contact with exposed portions of the conductor between the heater and the coolers, indicated a rise of temperature different according as the current was passing from hot to cold or from cold to hot. The heat seems therefore to be carried along by the current, and the process has accordingly been called the *electrical convection of heat.* In copper the heat moves with the current, in iron against it. In another form of statement, it may be said that, in unequally heated copper, a current from hot to cold heats the metal, and from cold to hot cools it, while in iron the reverse thermal effects occur. The experiments of Le Roux show that the process of electrical convection of heat cannot be detected in lead. For this reason, lead is used as the standard metal in constructing the thermo-electric diagram.

CHAPTER VII.

LUMINOUS EFECTS OF THE CURRENT.

237. **The Electric Arc.**—If the terminals of an electric circuit in which is an electromotive force of forty or more volts be formed of carbon rods, a brilliant and permanent luminous arc will appear between the ends of the rods if they be touched together and then withdrawn a short distance from each other. The temperature of the arc is so high that the most refractory substances melt or are dissipated when placed in it. The carbon forming the positive terminal is hotter than the other. Both the carbons are gradually oxidized, the loss of the positive terminal being about twice as great as that of the negative. The arc is, however, not due to combustion, since it can be formed in a vacuum.

The current passing in the arc is, in ordinary cases, not greater than ten amperes, while the measurements of the resistance of the arc show that it is altogether too small to account for this current when the original electromotive force is taken into account. This fact has been explained by Edlund and others on the hypothesis that there is a counter electromotive force set up in the arc, which diminishes the effective electromotive force of the circuit. The measurements of Lang show that this counter electromotive force in an arc formed between carbon points is about thirty-six volts, and in one formed between metal points about twenty-three volts.

238. **The Spark, Brush, and Glow Discharges.**—When a conductor is charged to a high potential and brought near another conductor which is joined to ground, a *spark* or a series

of sparks will pass from one to the other. This phenomenon and others associated with it are most readily studied by the use of an electrical machine or an induction coil, between the electrodes of which a great difference of potential can be easily produced. If the spark be examined with the spectroscope, its spectrum is found to be characterized by lines which are due to the metals composing the electrodes, and to the medium between them.

The passage of the spark through air or any dielectric is attended with a sharp report, and if the dielectric be solid, it is perforated or ruptured. If the electrodes be separated by a considerable distance, the path of the spark is usually a zigzag one. It is probable that this is due to irregularities in the dielectric, due to the presence of dust particles.

With proper adjustment of the electrodes, the discharge may sometimes be made to take the form of a long *brush* springing from the positive electrode, with a single trunk which branches and becomes invisible before reaching the negative electrode. Accompanying this is usually a number of small and irregular brushes starting from the negative electrode.

Another form of discharge consists of a pale luminous *glow* covering part of the surface of one or both electrodes. If a small conducting body be interposed between the electrodes when the glow is established, a portion of the glow will be cut off, marking out a region on the electrode which is the projection of the intervening conductor by the lines of electrical force. This phenomenon is called the *electrical shadow*.

The difference of potential required to set up a spark between two slightly convex metallic surfaces, separated by a stratum of air 0.125 centimetres thick, has been shown by Thomson to be about 5500 volts. The difference of potential which produces the sparks between the electrodes of an electrical machine, which are sometimes fifty or sixty centimetres long, must therefore be very great. The quantity of electricity which passes during the discharge is, however, exceedingly small, on account of the great resistance of the medium through which the discharge takes place.

Faraday showed that many of the phenomena of the discharge depend to some extent upon the medium in which it occurs. In chlorine the action seems to be the reverse of that in air, and those peculiar discharges which in air appear at the positive electrode appear in chlorine at the negative electrode.

It was proved by Franklin that the lightning flash is an electrical discharge between a cloud and the earth or another cloud at a different electrical potential. The differences of potential to which such discharges are due must be enormous, and the heat developed by the discharge shows that the quantity of electricity which passes in it is not inconsiderable.

Slowly moving fire-balls are sometimes seen, which last for a considerable time and disappear with a loud report and with all the attendant phenomena of a lightning discharge. It is not improbable that they are glow discharges which appear just before the difference of potential between the cloud and the earth becomes sufficiently great to give rise to a lightning flash.

239. **The Electrical Discharge in Rarefied Gases.**—If the air between the electrodes of an electrical machine be heated, it is found that the discharge takes place with greater facility and that the spark which can be obtained is longer than before. Similar phenomena appear if the air about the electrodes be rarefied by means of an air-pump. After the rarefaction has reached a certain point the discharge ceases to pass as a spark and becomes continuous. The arrangement in which this discharge is studied consists of a glass tube into which are sealed two platinum or, preferably, aluminium wires to serve as electrodes, and from which the air is removed to any required degree of exhaustion by an air-pump. Such an arrangement is usually called a *vacuum-tube.*

As the exhaustion proceeds there appears about the negative electrode in the tube a bright glow, separated from the electrode by a small non-luminous region. The body of the tube is filled with a faint rosy light, which in many cases breaks up into a succession of bright and dark layers transverse to the direction of the discharge. The discharge in this case is called the stratified discharge. A vacuum-tube in which the exhaustion is such that the phenomena are those here described is often called a *Geissler tube*. As the exhaustion is raised still higher, the rosy light in the tube fades out, the non-luminous space around the negative electrode becomes very much greater, and the phenomena in the tube become exceedingly interesting. They were discovered and have been carefully studied by Crookes, and the vacuum-tubes in which they appear are hence called *Crookes tubes*. They may be most conveniently described by assuming that there is a special discharge from the negative electrode, which we will usually call the discharge. This view receives some support from the fact that the relations of current and resistance in the tube are such as to indicate a counter electromotive force at the negative electrode.

The region occupied by the discharge from the negative electrode may be recognized by a faint blue light, which was not visible in the former condition of the tube. At every point on the wall of the tube to which this discharge extends occurs a brilliant phosphorescent glow, the color of which depends on the nature of the glass. The discharge seems to be independent of the position of the positive electrode, and to take place in nearly straight lines, which start normally from the negative electrode. If two negative electrodes be fixed in the tube, the discharge from one seems to be deflected by the other, and two discharges which meet at right angles seem to deflect one another.

If the discharge from a flat electrode be made to fall upon a body which can be moved, such as a glass film, or the vane of a light wheel, mechanical motions will be set up.

If the negative electrode be made in the form of a spherical cup, and a strip of platinum foil be placed at its centre, the foil will become heated to redness when the discharge is set up.

Two discharges in the same direction repel one another as if they were similarly electrified, and a magnet, brought near the outside of the tube, will deflect a discharge as if it were an electrical current.

The explanation of these phenomena is probably that given by Crookes, and adopted by Spottiswoode and Moulton. It is assumed that they are due to the presence of the molecules of gas left in the tube after the exhaustion has been brought to an end. The mean free path of the molecules in the tube is much greater than that at ordinary densities, and they can accordingly move through long distances in the tube before their original motion is checked by collision with other molecules. It is assumed that the molecules of gas in the tube are attracted by the negative electrode, are charged negatively by it, and are then repelled. The phenomena which have been described are then due to their collision with other bodies or with the wall of the tube, or to their mutual electrical repulsions and to the action between a moving quantity of electricity and a magnet.

The experiments of Spottiswoode and Moulton, who showed that the same phenomena appeared at lower exhaustions, if the intensity of the discharge were increased, are in favor of this explanation. So is also the fact that the Crookes phenomena appear with a maximum intensity at a certain period during the exhaustion of the tube, while if the exhaustion be carried as far as possible, by the help of chemical means, they cease altogether and no current passes in the tube. The connection of these phenomena with the action of the radiometer (§156) is also at once apparent.

SOUND.

CHAPTER I.

ORIGIN AND TRANSMISSION OF SOUND.

240. Definitions.—*Acoustics* has for its object the study of those phenomena which may be perceived by the ear. The sensations produced through the ear, and the causes that give rise to them, are called *sounds*.

241. Origin of Sound.—Sound is produced by vibratory movements in elastic bodies. The vibratory motion of bodies when producing sound is often evident to the eye. In some cases the sound seems to result from a continuous movement, but even in these cases the vibratory motion can be shown by means of an apparatus known as a *manometric* capsule, devised by König. It consists of a block A, Fig. 76, in which is a cavity covered by a membrane b. By means of a tube c illuminating gas is led into the cavity, and, passing out through the tube d, burns in a jet at e. It is evident that, if the membrane b be made to move suddenly inward or outward, it will compress or rarefy the gas in the capsule, and so cause the flow at the orifice and the height of the flame to increase or diminish. Any sound of sufficient intensity in the vicinity of the capsule causes an alternate lengthening and shortening of the flame, which, however, occur too frequently to be directly observed. By mov-

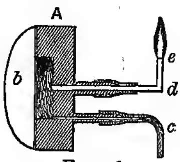

FIG. 76.

ing the eyes while keeping the flame in view, or by observing the image of the flame in a mirror which is turned from side to side, while the flame is quiescent, it appears drawn out into a broad band of light, but when it is agitated by a sound near it, it appears serrate on its upper edge or even as a series of separate flames. This lengthening and shortening of the flame is evidence of a to-and-fro movement of the membrane, and hence of the sounding body that gave rise to the movement. If a hole be made in the side of an organ-pipe and the capsule made to cover it, the vibrations of the air-column within the pipe may be shown. By suitable devices the vibratory motion of all sounding bodies may be demonstrated.

242. Propagation of Sound.—The vibratory motion of a sounding body is ordinarily transmitted to the ear through the air. This is proved by placing a sounding body under the receiver of an air-pump and exhausting the air. The sound becomes fainter and fainter as the exhaustion proceeds, and finally becomes inaudible if the vacuum is good. Sound may, however, be transmitted by any elastic body.

In order to study the character of the motion by which sound is propagated, let us suppose *AB* (Fig. 77) to represent

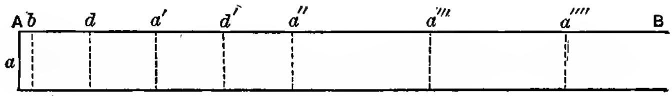

Fig. 77.

a cylinder of some elastic substance, and suppose the layer of particles *a* to suffer a small displacement to the right. The effect of this displacement is not immediately to move forward the succeeding layers, but *a* approaches *b*, producing a condensation, and developing a force that soon moves *b* forward; this in turn moves forward the next layer, and so the motion is transmitted from layer to layer through the cylinder with a velocity that depends upon the elasticity (§ 76) of the substance, and upon its density. This velocity is expressed by the formula $V=\sqrt{\frac{E}{D}}$, in which E represents the elasticity of the substance, and D its density (§ 268). Now, if we suppose the layer a, from any cause whatever, to execute regular vibrations, this movement will be transmitted to the succeeding layers with the velocity given by the formula, and, in time, each layer of particles in the cylinder will be executing vibrations similar to those of a. If the vibrations of a be performed in the time t, the motion will be transmitted during one complete vibration of a to a distance $s = vt$, where v is the velocity of propagation, say to a', during two complete vibrations of a, to a distance $2s = 2vt$, or to a'', during three complete vibrations to a''', and so on. It is evident that the layer a' begins its first vibration at the instant that a begins its second vibration, a'' begins its first vibration at the instant that a' begins its second, and a its third vibration. The layer midway between a and a' evidently begins its vibration just as a completes the first half of its vibration, and therefore moves forward while a moves backward. This condition of things existing in the cylinder constitutes a *wave motion.* While a moves forward, the portions near it are compressed. While it moves backward, they are dilated. Whatever the condition at a, the same condition will exist at the same instant at a', a'', etc. The distance $aa' = a'a''$ is called a *wave length;* it is the distance from any one particle to the next one of which the vibrations are in the same phase (§ 19). If the condition at a and a' be one of condensation, it is evident that at d, midway between a and a', there must be a rarefaction. In the wave length $aa_{\prime}$ exist all intermediate conditions of condensation and rarefaction. These conditions must follow each other along the cylinder with the velocity of the transmitted motion, and they constitute a *progressive wave* moving with this velocity. If the vibratory motion with which *a* is endowed be communicated by a sounding body, the wave is a *sound-wave.* If, instead of a cylinder of the substance, we have an indefinite medium in the midst of which the sounding body is placed, the motion is transmitted in all directions as spherical waves about the sounding body as a centre.

243. Mode of Propagation of Wave Motion.—The mode of transmission of wave motions was first shown by Huyghens, and the principle involved is known as *Huyghens' principle.* Let *a* (Fig. 78) be a centre from which sound originates. At the end of a certain time it will have reached the surface *mn.* From the preceding discussion it is evident that each particle of the surface *mn* has a vibratory motion similar to that at *a.* Any one of those particles would, if vibrating alone, be, like *a,* the centre of a system of spherical waves, and each of them must, therefore, be considered as a wave centre from which spherical waves proceed. Suppose such a wave to proceed from each one of them for the short distance *cd.* Since the number of the elementary spherical waves is very great, it is plain that they will coalesce to form the surface *m'n'* which determines a new position of the wave surface. In some cases the existence of these elementary waves need not be considered, but there are many phenomena of wave motion which can only be studied by recognizing the fact that propagation always takes place as above described.

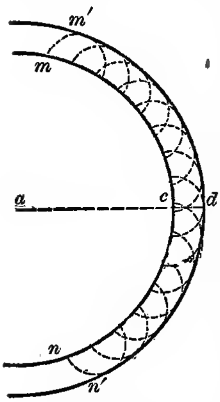

FIG. 78.

244. Graphic Representation of Wave Motion.—In order to study the movements of a body in which a wave motion exists, especially when two or more systems of waves exist in the same body, it is convenient to represent the movement by a sinusoidal curve, as described in § 19.

Suppose the layer a (Fig. 77) to move with a simple harmonic motion of which the amplitude is a and the period T, and let time be reckoned from the instant that the particles pass the position of equilibrium in a positive direction. A sinusoidal curve may be constructed to represent either the displacements of the various layers from their positions of equilibrium, or the velocities with which they are severally moving at a given time.

To construct the first curve let the several points along OX (Fig. 79) represent points of the body through which the wave

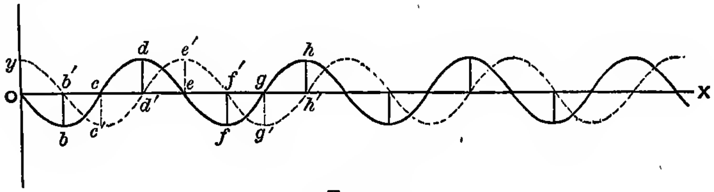

FIG. 79.

is moving. Let $Oy = a$ be the amplitude of vibration of each particle. The displacement of the particle at O at any instant t after passing its position of equilibrium is $y = a \cos\left(\frac{2\pi t}{T} - \frac{\pi}{2}\right)$, since when t is reckoned from the position of equilibrium $\epsilon = \frac{\pi}{2}$. Hence $y = a \sin \frac{2\pi t}{T}$. If v represent the velocity of propagation of the wave, the particle at the distance x from the origin will have a displacement equal to that of the particle at O at the instant t, at an instant later than t by the time taken for the wave to travel over the distance x, or $\frac{x}{v}$ seconds. Hence its displacement at the instant t will be the same as that which existed at O, $\frac{x}{v}$ seconds earlier. But the displacement at O, $\frac{x}{v}$ seconds earlier is

$$y = a \sin \frac{2\pi\left(t - \frac{x}{v}\right)}{T}$$

$$= a \sin 2\pi\left(\frac{t}{T} - \frac{x}{vT}\right). \qquad (102)$$

The quantity vT equals the distance through which the movement is transmitted during the time of one complete vibration of the particle at O. Putting this equal to λ, we have finally

$$y = a \sin 2\pi\left(\frac{t}{T} - \frac{x}{\lambda}\right). \qquad (103)$$

Suppose $t = 0$, and give to x various values. The corresponding values of y will represent the displacement at that instant of the particle the distance of which from the origin is x. For $x = 0$, $y = 0$. For $x = \frac{1}{4}\lambda$, $y = -a$. For $x = \frac{1}{2}\lambda$, $y = 0$. For $x = \frac{3}{4}\lambda$, $y = a$. For $x = \lambda$, $y = 0$, etc. Laying off these values of x on OX and erecting perpendiculars equal to the corresponding values of y, we have the curve *Obcde* . . . .

The above expression for y may be put in the form

$$y = a \sin 2\pi\left(\frac{\frac{t\lambda}{T} - x}{\lambda}\right).$$

Hence, if any finite value be assigned to t, we shall obtain for y the same values as were obtained above for $t = 0$, if we increase each of the values of x by $\frac{t\lambda}{T}$. For instance, if t equal

$\frac{1}{4}T$, we have $y = 0$ for $x = \frac{1}{4}\lambda$, $y = -a$ for $x = \frac{1}{2}\lambda$, etc., and the curve becomes the dotted line $b'c'd'$ . . . . The effect of increasing t is to displace the curve along OX in the direction of propagation of the wave.

The formula for constructing the curve of velocities is derived in the same way as that for displacements. It is

$$y = \frac{2\pi a}{T} \cos 2\pi \left(\frac{t}{T} - \frac{x}{\lambda}\right). \tag{104}$$

Fig. 80 shows the relation of the two curves. The upper is the curve of displacement, and the lower of velocity.

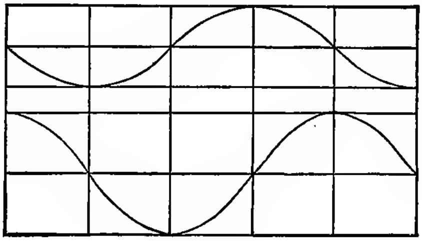

FIG. 80.

245. **Composition of Wave Motions.**—The composition of wave motions may be studied by the help of the curves explained above. If two systems of waves coexist in the same body, the displacement of any particle at any instant will be the algebraic sum of the displacements due to the systems taken separately. If the curve of displacements be drawn for each system, the algebraic sum of the ordinates will give the ordinates of the curve representing the actual displacements. In

Fig. 81 the dotted line and the light full line represent respectively the displacements due to two wave systems of the same period and amplitude. The heavy line represents the actual displacement. In I the two systems are in the same phase; in II the phases differ by $\frac{1}{4}$, and in III by $\frac{1}{2}$, of a period. If both wave systems move in the same direction, it is evident that the conditions of the body will be continuously shown by supposing the heavy line to move in the same direction with the same velocity. The condition represented in III is of special interest. It shows that two wave systems may completely annul

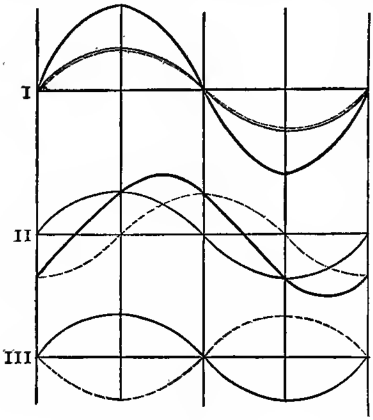

FIG. 81.

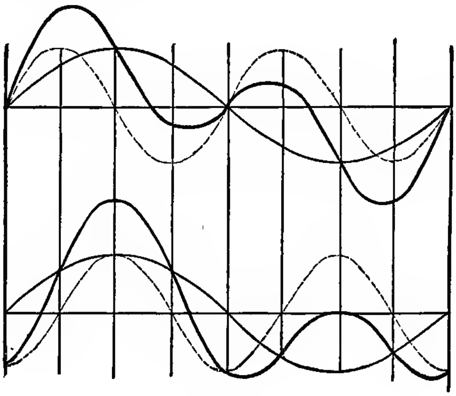

FIG. 82.

each other. Fig. 82 represents the resultant wave when the periods, and consequently the wave lengths, of the two systems are as 1 : 2. It will be noticed that the resultant curve is no longer a simple sinusoid.

In the same way the resultant wave may be constructed for any number of wave systems having any relation of wave lengths, amplitudes, and phases. A very important case is that of two wave systems of the same period moving in opposite directions with the same velocity. In this case the two systems no longer maintain the same relative positions, and the resultant curve is not displaced along the axis, but continually changes form. In Fig. 83, let the full and dotted lines in I represent, at a given instant, the displacements due to the two waves respectively. The resultant is plainly the straight line *ab*, which indicates that at that instant there is no displacement of any particle. At an instant later by $\frac{1}{8}$ period, as shown in II, the wave represented by the full line has moved to the right $\frac{1}{8}$ wave length, while that represented by the dotted line has moved to the left the same distance. The heavy line indicates the corresponding displacements. In III, IV, V, etc., the conditions at instants $\frac{1}{4}$, $\frac{3}{8}$, $\frac{1}{2}$, etc., periods later are represented. A comparison of these

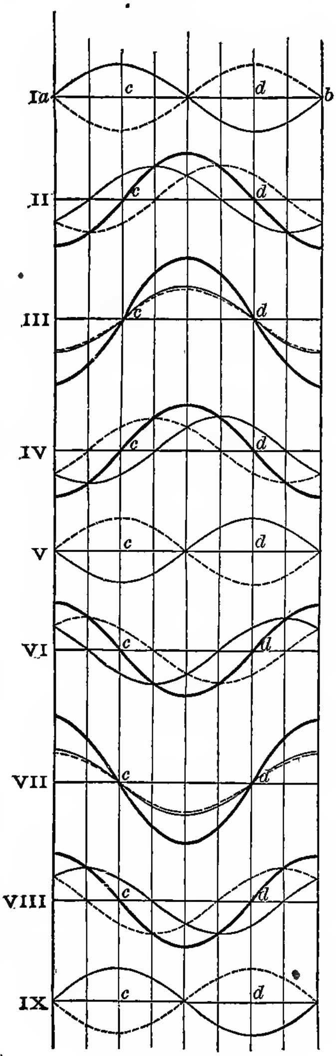

FIG. 83.

figures will show that the particles at *c* and *d* are always at rest, that the particles between *c* and *d* all move in the same direction at the same time, and that particles on the opposite sides of *c* or *d* are always moving in opposite directions. It follows that the resultant wave has no progressive motion. It is a *stationary wave.* Places where no motion occurs, such as *c* and *d*, are called *nodes.* The space between two nodes is an *internode* or *ventral segment.* The middle of a ventral segment, where the motion is greatest, is an *anti-node.* It will be seen later that all sounding bodies afford examples of stationary waves.

246. Reflection of Waves.—When a wave reaches the bounding surface between two media, one of three cases may occur:

(1) The particles of the second medium may have the same facility for movement as those of the first. The condition at the boundary will then be the same as that at any point previously traversed, and the wave will proceed as though the first medium were continuous.

(2) The particles of the second medium may move with less facility than those of the first. Then the condensed portion of a wave which reaches the boundary becomes more condensed in consequence of the restricted forward movement of the bounding particles, and the rarefied portion becomes more rarefied, because those particles are also restricted in their backward motion. The condensation and rarefaction are communicated backward from particle to particle of the first medium, and constitute a *reflected wave.* It will be seen that, when the condensed portion of the wave, in which the particles have a forward movement, reaches the boundary, the effect is a greater condensation, that is, the same effect as would be produced by imparting a backward movement to the bounding particles if no wave previously existed. In the direct rarefied portion of the wave the movement of the particles is backward, and the effect, at the boundary, of a greater rarefaction is what would

be produced by a forward movement of those particles. The effect in this case is, therefore, to reverse the motion of the particles. It is called *reflection with change of sign.*

(3) The particles of the second medium may move more freely than those of the first. In this case, when a wave in the first medium reaches the boundary, the bounding particles, instead of stopping with a displacement such as they would reach in the interior of the medium, move to a greater distance, and this movement is communicated back from particle to particle as a reflected wave in which the motion has the same sign as in the direct wave. It is *reflection without change of sign.* The two latter cases are extremely important in the study of the formation of stationary waves in sounding bodies.

247. Law of Reflection.—Let us suppose a system of spherical waves departing from the point C (Fig. 84). Let mn be the intersection of one of the waves with the plane of the paper. Let AB be the trace of a plane smooth surface perpendicular to the plane of the paper, upon which the waves impinge. mo shows the position which the wave of which mn is a part would have occupied had it not been intercepted by the surface. From the last section it appears that reflection will take place as the wave mno

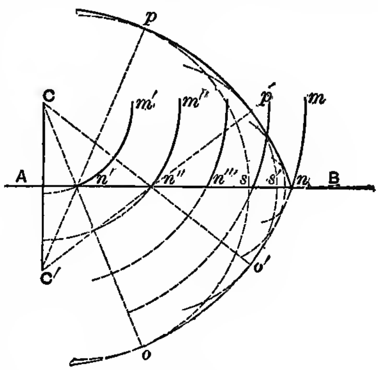

FIG. 84.

strikes the various points of AB. In § 243 it was seen that any point of a wave may be considered as the centre of a wave system, and we may therefore take n', n'', etc., the points of intersection of the surface AB with the wave mn when it occupied the positions $m'n'$, $m''n''$, etc., as the centres of systems of spherical waves, the resultant of which would be the actual wave proceeding from AB. With n' as a centre describe a sphere tangent to mno at o. It is evident that this will represent the elementary spherical wave of which the centre is n' when the main wave is at mn. Describe similar spheres with n'', n''', etc., as centres. The surface np, which envelops and is tangent to all these spheres, represents the wave reflected from AB. If that part of the plane of the paper below AB be revolved about AB as an axis until it concides with the paper above AB, so will coincide with sp, $s'o'$ with $s'p'$, etc., and hence no with np. But no is a circle with C as a centre; np is, therefore, a circle of which the centre is C', on a perpendicular to AB through C, and as far below AB as C is above. When, therefore, a wave is reflected at a plane surface, the centres of the incident and reflected waves are on the same line perpendicular to the reflecting surface, and at equal distances from the surface on opposite sides.

CHAPTER II.

SOUNDS AND MUSIC.

COMPARISON OF SOUNDS.

248. Musical Tones and Noises.—The distinction between the impressions produced by musical tones and by noises is familiar to all. Physically, a musical *tone* is a sound the vibrations of which are regular and periodic. A *noise* is a sound the vibrations of which are very irregular. It may result from a confusion of musical tones, and is not always devoid of musical value. The sound produced by a block of wood dropped on the floor would not be called a musical tone, but if blocks of wood of proper shape and size be dropped upon the floor in succession, they will give the tones of the musical scale.

Musical tones may differ from one another in *pitch*, depending upon the frequency of the vibrations; in *loudness*, depending upon the amplitude of vibration; and in *quality*, depending upon the manner in which the vibration is executed. In regard to pitch, tones are distinguished as *high* or *low*, *acute* or *grave*. In regard to loudness, they are distinguished as *loud* or *soft*. The quality of musical tones enables us to distinguish the tones of different instruments even when sounding the same notes.

249. Methods of Determining the Number of Vibrations of a Musical Tone.—That the pitch of a tone depends upon the frequency of vibrations may be simply shown by holding the corner of a card against the teeth of a revolving wheel. With a very slow motion the card snaps from tooth to tooth, making a succession of distinct taps, which, when the revolutions are sufficiently rapid, blend together and produce a continuous tone, the pitch of which rises and falls with the changes of speed. Savart made use of such a wheel to determine the number of vibrations corresponding to a tone of given pitch. After regulating the speed of rotation until the given pitch was reached, the number of revolutions per second was determined by a simple attachment; this number multiplied by the number of teeth in the wheel gave the number of vibrations per second.

The *siren* is an instrument for producing musical tones by puffs of air succeeding each other at short equal intervals. A circular disk having in it a series of equidistant holes arranged in a circle around its axis is supported so as to revolve parallel to and almost touching a metal plate in which is a similar series of holes. The plate forms one side of a small chamber, to which air is supplied from an organ bellows. If there be twenty holes in the disk, and if it be placed so that these holes correspond to those in the plate, air will escape through all of them. If the disk be turned through a small angle, the holes in the plate will be covered and the escape of air will cease. If the disk be turned still further, at one twentieth of a revolution from its first position, air will again escape, and if it rotate continuously, air will escape twenty times in a revolution. When the rotation is sufficiently rapid, a continuous tone is produced the pitch of which rises as the speed increases. The siren may be used exactly as the toothed wheel to determine the number of vibrations corresponding to any tone.

By drilling the holes in the plate obliquely forward in the direction of rotation, and those in the disk obliquely backward, the escaping air will cause the disk to rotate, and the speed of rotation may be controlled by controlling the pressure of air in the chamber.

Sirens are sometimes made with several series of holes in the disk. These serve not only the purposes described above, ıut also to compare tones of which the vibration numbers have ertain ratios.

The number of vibrations of a ounding body may sometimes be deermined by attaching to it a light tylus which is made to trace a curve .pon a smoked glass or cylinder. Intead of attaching the stylus to the ounding body directly, which is pracicable only in a few cases, it may be atached to a membrane which is caused o vibrate by the sound-waves which he body generates. A membrane reroduces very faithfully all the characeristics of the sound-waves, and the urve traced by the stylus attached to : gives information, therefore, not nly in regard to the number of vibraions, but to some extent in regard ɔ their amplitude and form.

PHYSICAL THEORY OF MUSIC.

250. Concord and Discord.— Vhen two or more tones are sounded ɔgether, if the effect be pleasing there said to be *concord;* if harsh, *discord.* 'o understand the cause of discord, ıppose two tones of nearly the same itch to be sounded together. The reıltant curve, constructed as in § 245, like those in Fig. 85, which repreːnt the resultants when the periods ⁚ the components have the ratio 81 : 80 and when they have

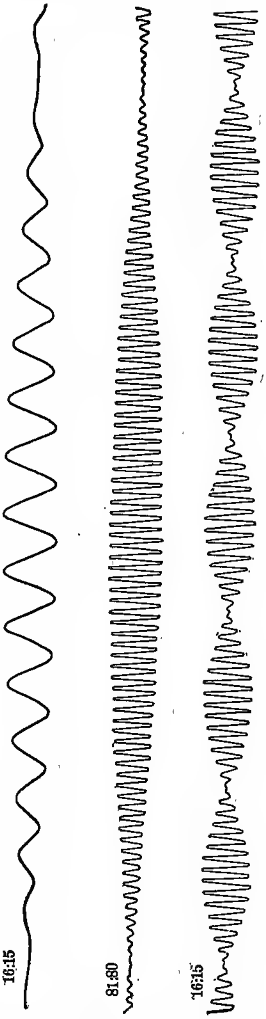

FIG. 85.

the ratio 16 : 15. The figure indicates, what experiment verifies, that the resultant sound suffers periodic variations in intensity. When these variations occur at such intervals as to be readily distinguished, they are called *beats*. These beats occur more and more frequently as the numbers expressing the ratio of the vibrations reduced to its lowest terms become smaller, until they are no longer distinguishable as separate beats, but appear as an unpleasant roughness in the sound. If the terms of the ratio become smaller still, the roughness diminishes, and when the ratio is $\frac{6}{5}$ the effect is no longer unpleasant. This, and ratios expressed by smaller numbers, as $\frac{5}{4}$, $\frac{5}{3}$, $\frac{4}{3}$, $\frac{3}{2}$, $\frac{2}{1}$, represent concordant combinations.

251. Major and Minor Triads.—Three tones of which the vibration numbers are as 4 : 5 : 6 form a concordant combination called the *major triad*. The ratio 10 : 12 : 15 represents another concordant combination called the *minor triad*. Fig. 86 shows the resultant curves for the two triads.

FIG. 86.

252. Intervals.—The *interval* between two tones is expressed by the ratio of their vibration numbers, using the larger as the numerator. Certain intervals have received names derived from the relative positions of the two tones in the musical scale, as described below. The interval $\frac{2}{1}$ is called an *octave*; $\frac{3}{2}$, a *fifth*; $\frac{4}{3}$, a *fourth*; $\frac{5}{4}$, a *major third*; $\frac{6}{5}$, a *minor third*.

253. Musical Scales.—A *musical scale* is a series of tones which have been chosen to meet the demands of musical composition. There are at present two principal scales in use, each consisting of seven notes, with their octaves, chosen with reference to their fitness to produce pleasing effects when used in combination. In one, called the *major scale*, the first, third, and fifth, the fourth, sixth, and eighth, and the fifth, seventh, and ninth tones, form major triads. In the other, called the *minor scale*, the same tones form minor triads. From this it is easy to deduce the following relations:

MAJOR SCALE.

								1′	2′
Tone Number..............	1	2	3	4	5	6	7	8	9
Letter......................	C	D	E	F	G	A	B	C′	D′
Name......................	do or ut	re	mi	fa	sol	la	si	ut	re
Number of vibrations........	m	$\frac{9}{8}$m	$\frac{5}{4}$m	$\frac{4}{3}$m	$\frac{3}{2}$m	$\frac{5}{3}$m	$\frac{15}{8}$m	2m	$\frac{9}{4}$m
Intervals from tone to tone..		$\frac{9}{8}$	$\frac{10}{9}$	$\frac{16}{15}$	$\frac{9}{8}$	$\frac{10}{9}$	$\frac{9}{8}$	$\frac{16}{15}$	

MINOR SCALE.

Tone Number..............	1	2	3	4	5	6	7	8	9
Letter......................	A	B	C	D	E	F	G	A′	B′
Name......................	la	si	ut	re	mi	fa	sol	la	si
Number of vibrations........	m	$\frac{9}{8}$m	$\frac{6}{5}$m	$\frac{4}{3}$m	$\frac{3}{2}$m	$\frac{8}{5}$m	$\frac{9}{5}$m	2m	$\frac{9}{4}$m
Intervals from tone to tone..		$\frac{9}{8}$	$\frac{16}{15}$	$\frac{10}{9}$	$\frac{9}{8}$	$\frac{16}{15}$	$\frac{9}{8}$	$\frac{10}{9}$	

The derivation of the names of the intervals will now be apparent. For example, an interval of a third is the interval between any tone of the scale and the third one from it, counting the first as 1. If we consider the intervals from tone to tone, it is seen that the pitch does not rise by equal steps, but that there are three different intervals, $\frac{9}{8}$, $\frac{10}{9}$, and $\frac{16}{15}$. The first two are usually considered the same, and are called *whole tones*. The third is a *half-tone* or *semitone*.

It is desirable to be able to use any tone of a musical instrument as the first tone or *tonic* of a musical scale. To permit this, when the tones of the instrument are fixed, it is plain that extra tones, other than those of the simple scale, must be provided in order that the proper sequence of intervals may be maintained. Suppose the tonic to be transposed from C to D.

The semitones should now come, in the major scale, between F and G, and C′ and D′, instead of between E and F, and B and C′. To accomplish this, a tone must be substituted for F and another for C′. These are called F sharp and C′ sharp respectively, and their vibration numbers are determined by multiplying the vibration numbers of the tones which they replace by $\frac{25}{24}$. The introduction of five such extra tones, making twelve in the octave, enables us to preserve the proper sequence of whole tones and semitones, whatever tone is taken as the tonic. But if we consider that the whole tones are not all the same, and propose to preserve exactly all the intervals of the transposed scale, the problem becomes much more difficult, and can only be solved at the expense of too great complication in the instrument. Instead of attempting it, a system of tuning, called *temperament*, is used by which the twelve tones referred to above are made to serve for the several scales, so that while none are perfect, the imperfections are nowhere marked. The system of temperament usually employed, or at least aimed at, called the *even temperament*, divides the octave into twelve equal semitones, and each interval is therefore the twelfth root of 2. With instruments in which the tones are not fixed, like the violin for instance, the skilful performer may give them their exact value.

For convenience in the practice of music and in the construction of musical instruments, a *standard pitch* must be adopted. This pitch is usually determined by assigning a fixed vibration number to the tone above the middle C of the piano, represented by the letter A′. This number is about 440, but varies somewhat in different countries and at different times. In the instruments made by König for scientific purposes, the vibration number 256 is assigned to the middle C. This has the advantage that the vibration numbers of the successive octaves of this tone are powers of 2.

CHAPTER III.

VIBRATIONS OF SOUNDING BODIES.

254. **General Considerations.**—The principles developed in § 246 apply directly in the study of the vibrations of sounding bodies. When any part of a body which is capable of acting as a sounding body is set in vibration, a wave is propagated through it to its boundaries, and is there reflected. The reflected wave, travelling away from the boundary, in conjunction with the direct wave going toward it, produces a stationary wave. These stationary waves are characteristic of the motion of all sounding bodies. Fixed points of a body often determine the position of nodes, and in all cases the length of the wave must have some relation to the dimensions of the body.

255. **Organ Pipes.**—A column of air, enclosed in a tube of suitable dimensions, may be made to vibrate and become a sounding body. Let us suppose a tube closed at one end and open at the other. If the air particles at the open end be suddenly moved inward, a pulse travels to the closed end, and is there reflected with change of sign (§ 246). It returns to the open end and is again reflected, this time without change of sign, because there is greater freedom of motion without than within the tube. As it starts again toward the closed end, the air particles that compose it move outward instead of inward. If they now receive an independent impulse outward, the two effects are added and a greater disturbance results. So, by properly timing small impulses at the open end of the tube, the air in it may be made to vibrate strongly.

If a continuous vibration be maintained at the open end of the tube, waves follow each other up the tube, are reflected with change of sign at the closed end, and returning, are reflected without change of sign at the open end. Any given wave *a*, therefore, starts up the tube the second time with its phase changed by half a period. The direct wave that starts up the tube at the same instant must be in the same phase as the reflected wave, and it therefore differs in phase half a period from the direct wave *a*. In other words, any wave returning to the mouth-piece must find the vibrations there opposite in phase to those which existed when it left. This is possible only when the vibrating body makes, during the time the wave is going up the tube and back, 1, 3, 5, or some odd number of half-vibrations. By constructing the curves representing the stationary wave resulting from the superposition of the two systems of vibrations, it will be seen that there is always a node at the closed end of the tube and an anti-node at the mouth. When there is 1 half-vibration while the wave travels up and back, the length of the tube is $\frac{1}{4}$ the wave length; when there are 3 half-vibrations in the same time, the length of the tube is $\frac{3}{4}$ the wave length, and there is a node at one third the length of the tube from the mouth.

If the tube be open at both ends, reflection without change of sign takes place in both cases, and the reflected wave starts up the tube the second time in the same phase as at first. The vibrations must therefore be so timed that 1, 2, 3, 4, or some whole number of complete vibrations are performed while the wave travels up the tube and back. A construction of the curve representing the stationary wave in this case will show, for the smallest number of vibrations, a node in the middle of the tube and an anti-node at each end. The length of the tube is therefore $\frac{1}{2}$ the wave length for this rate of vibration. The vibration numbers of the several tones produced by an open tube are evidently in the ratio of the series of whole numbers 1, 2, 3, 4, etc., while for the closed tube only those tones can be produced of which the vibration numbers are in the ratio of the series of odd numbers 1, 3, 5, etc. It is evident also that the lowest tone of the closed tube is an octave lower than that of the open tube.

This lowest tone of the tube is called the *fundamental*, and the others are called *overtones*, or *harmonics*. These simple relations between the length of the tube and length of the wave are only realized when the tubes are so narrow that the air particles lying in a plane cross-section are all actuated by the same movement. This is never the case at the open end of the tube, and the distance from this end to the first node is, therefore, always less than a quarter wave length.

256. Modes of Exciting Vibrations in Tubes.—If a tuning fork be held in front of the open mouth of a tube of proper length, the sound of the fork is strongly reinforced by the vibration of the air in the tube. If we merely blow across the open end of a tube, the agitation of the air may, by the reaction of the returning reflected pulses, be made to assume a regular vibration of the proper rate and the column made to sound. In organ pipes a mouthpiece of the form shown in Fig. 87 is often employed. The thin sheet of air projected against the thin edge is thrown into vibration. Those elements of this vibration which correspond in frequency with the pitch of the pipe are strongly reinforced by the action of the stationary wave set up in the pipe, and hence the tone proper to the pipe is produced. Sometimes *reeds* are used, as shown in Fig. 87*a*. The air escaping from the chamber *a* through the passage *c* causes the reed *r* to vibrate. This alternately closes and opens the passage,

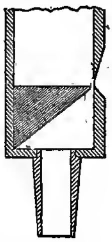

FIG. 87.

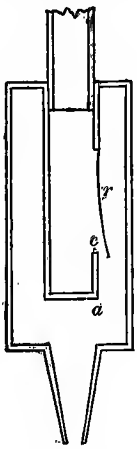

FIG. 87*a*.

and so throws into vibration the air in the pipe. If the reed be stiff, and have a determined period of vibration of its own, it must be tuned to suit the period of the air column which it is intended to set in vibration. If the reed be very flexible it will accommodate itself to the rate of vibration of the air column, and may then serve to produce various tones, as in the clarionet.

In instruments like the cornet and bugle, the lips of the player act as a reed, and the player may at will produce many of the different overtones. In that way melodies may be played without the use of keys or other devices for changing the length of the air column.

Vibrations may be excited in a tube by placing a gas flame at the proper point in it. The flame thus employed is called a *singing flame*. The organ of the voice is a kind of reed pipe in which little folds of membrane, called vocal chords, serve as reeds which can be tuned to different pitches by muscular effort, and the cavity of the mouth and larynx serves as a pipe in which the mass of air may also be changed at will, in form and volume.

257. Longitudinal Vibrations of Rods.—A rod free at both ends vibrates as the column of air in an open tube. Any displacement produced at one end is transmitted with the velocity of sound in the material to the other end, is there reflected without change of sign and returns to the starting point to be reflected again exactly as in the open tube. The fundamental tone corresponds to a stationary wave having a node at the centre of the rod.

258. Longitudinal Vibrations of Cords.—Cords fixed at both ends may be made to vibrate by rubbing them lengthwise. Here reflection with change of sign takes place at both ends, which brings the wave as it leaves the starting point the second time to the same phase as when it first left it, and there must be, therefore, as in the open tube, 1, 2, 3, 4, etc., vibrations while the wave travels twice the length of the cord. The velocity of transmission of a longitudinal displacement in a wire depends upon the elasticity and density of the material only. The velocity and the rate of vibration are, therefore, nearly independent of the stretching force.

259. Transverse Vibrations of Cords.—If a transverse vibration be given to a point upon a wire fastened at both ends, everything relating to the reflection of the wave motion and the formation of stationary waves is the same as for longitudinal displacements. The velocity of transmission, and consequently the frequency of the vibrations, are, however, very different. If the cord offer no resistance to flexure, the force tending to restore it to its position of equilibrium is entirely due to the stretching force. This, therefore, takes the place of the elasticity in the formula for transmission of longitudinal vibrations (§ 268). The mass of the cord per unit length takes the place of the density in the same formula. Thus we have the formula for the velocity

$$V = \sqrt{\frac{P}{m}},$$

where P is the stretching force and m the mass per unit length. The greatest time of vibration, the time required for the wave to travel twice the length of the string, is

$$T = \frac{2L}{V} = 2L\sqrt{\frac{m}{P}}, \qquad (105)$$

and the number of vibrations per second is

$$N = \frac{1}{T} = \frac{1}{2L}\sqrt{\frac{P}{m}}. \qquad (106)$$

Hence, the number of vibrations of a string is inversely as the length, directly as the square root of the tension, and inversely as the square root of the mass per unit length. These laws are readily verified by experiment.

260. Transverse Vibrations of Rods, Plates, etc.—The vibrations of rods, plates, and bells are all cases of stationary waves resulting from systems of waves travelling in opposite directions. Subdivision into segments occurs, but, in these cases, the relations of the various overtones are not so simple as in the cases before considered. For a rod fixed at one end, sounding its fundamental tone, there is a node at the fixed end only. For the first overtone there is a second node near the free end of the rod, and the number of vibrations is a little more than six times the number for the fundamental.

A rod free at both ends has two nodes when sounding its fundamental, as shown in Fig. 88. The distance of these nodes from the ends is about $\frac{2}{9}$ the length of the rod. If the rod be bent, the nodes approach the centre until, when it has assumed the U form like a tuning-fork, the two nodes are very near the centre. This will be understood from Fig. 89.

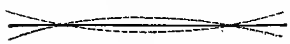

Fig. 88.

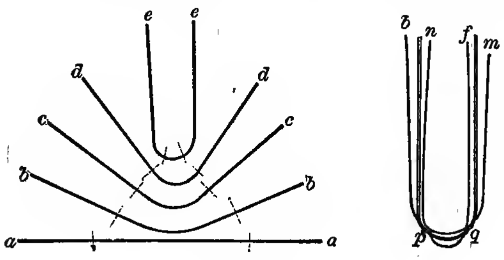

Fig. 89.

The nodal lines on plates may be shown by fixing the plate in a horizontal position and sprinkling sand over its surface. When the plate is made to vibrate, the sand gathers at the nodes and marks their position. The figures thus formed are known as Chladni's figures.

261. Communication of Vibrations.—If several pendulums be suspended from the same support, and one of them be made to vibrate, any others which have the same period of vibration will soon be found in motion, while those which have a different period will show no signs of disturbance. The vibration of the first pendulum produces a slight movement of the support which is communicated alike to all the other pendulums. Each movement may be considered as a slight impulse, which imparts to each pendulum a very small vibratory motion. For those pendulums having the same period as the one in vibration, these impulses come just in time to increase the motion already produced, and so, after a time, produce a sensible motion; while for those pendulums having a different period, the vibration at first imparted will not keep time with the impulses, and these will therefore as often tend to destroy as to increase the motion. It is important to note that the pendulum imparting the motion loses all it imparts. This is not only true of pendulums, but of all vibrating bodies. Two strings stretched from the same support and tuned to unison will both vibrate when either one is caused to sound. A tuning-fork suitably mounted on a sounding-box will communicate its vibrations to another tuned to exact unison even when they are thirty or forty feet apart and only air intervenes. In this case it is the sound-wave generated by the first fork which excites the second fork, and in so doing the wave loses a part of its own motion, so that beyond the second fork, on the line joining the two, the sound will be less intense than at the same distance in other directions.

Air columns of suitable dimensions will vibrate in sympathy with other sounding bodies. If water be gradually poured into a deep jar, over the mouth of which is a vibrating tuning-fork, there will be found in general a certain length of the air column for which the tone of the fork is strongly reinforced. From the theory of organ pipes, it is plain that this length corresponds approximately to a quarter wave length for that tone. In this case, also, when the strongest reinforcement occurs, the sound of the fork will rapidly die away. The sounding-boxes on which the tuning-forks made by König are mounted are of such dimensions that the enclosed body of air will vibrate in unison with the fork, but they are purposely made not quite of the dimensions for the best resonance, in order that the forks may not too quickly be brought to rest.

A membrane or a disk, fastened by its edges, may respond to and reproduce more or less faithfully a great variety of sounds. Hence such disks, or *diaphragms*, are used in instruments like the telephone and phonograph, designed to reproduce the sounds of the voice. The *phonograph* consists of a mouthpiece and disk similar to that used in the telephone, but the disk has fastened to its centre, on the side opposite the mouthpiece, a short stiff stylus, which serves to record the vibrations of the disk upon a sheet of tinfoil or wax moved along beneath it. The foil is wrapped upon a cylinder having a spiral groove on its surface, and upon its axle a screw thread of the same pitch works in a fixed nut so that, when the cylinder revolves, it has also an endwise motion, such that a fixed point would follow the spiral groove on its surface. To use the instrument, the disk is placed in position with the stylus attached adjusted to enter the groove in the cylinder and slightly indenting the foil. The cylinder is revolved while sounds are produced in front of the disk. The disk vibrates, causing the stylus to indent the foil more or less deeply, so leaving a permanent record. If now the cylinder be turned back to the starting-point and then turned forward, causing the stylus to go over again the same path, the indentations previously made in the foil now cause the stylus, and consequently the disk, to vibrate and reproduce the sound that produced the record.

The sounding-boards of the various stringed instruments are in effect thin disks, and afford examples of the reinforcement of vibrations of widely different pitch and quality by the same body. The strings of an instrument are of themselves insufficient to communicate to the air their vibrations, and it is only through the sounding-board that the vibrations of the string can give rise to audible sounds. The quality of stringed instruments, therefore, depends largely upon the character of the sounding-board.

The tympanum of the ear furnishes another example of the facility with which membranes respond to a great variety of sounds.

CHAPTER IV.

ANALYSIS OF SOUNDS AND SOUND SENSATIONS.

262. **Quality.**—As has already been stated, the tones of different instruments, although of the same pitch and intensity, are distinguished by their *quality*. It was also stated that the quality of a tone depends upon the manner in which the vibration is executed. The meaning of this statement can best be understood by considering the curves which represent the

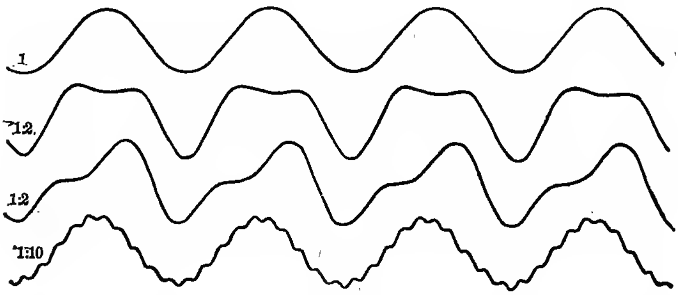

FIG. 90.

vibrations. In Fig. 90 are given several forms of vibration curves of the same period.

Every continuous musical tone must result from a periodic vibration, that is, a vibration which, however complicated it may be, repeats itself at least as frequently as do the vibrations of the lowest audible tone. According to Fourier's theorem (§ 19), every periodic vibration is resolvable into simple harmonic vibrations having commensurable periods. It has been seen that all sounding bodies may subdivide into segments, and produce a series of tones of which the vibration periods generally bear a simple relation to each other. These may be produced simultaneously by the same body, and so give rise to complex tones the character of which will vary with the nature and intensity of the simple tones produced. It has been held that the quality of a complex tone is not affected by change of phase of the component simple tones relative to each other. Some experiments by König seem to indicate, however, that the quality does change when there is merely change of phase.

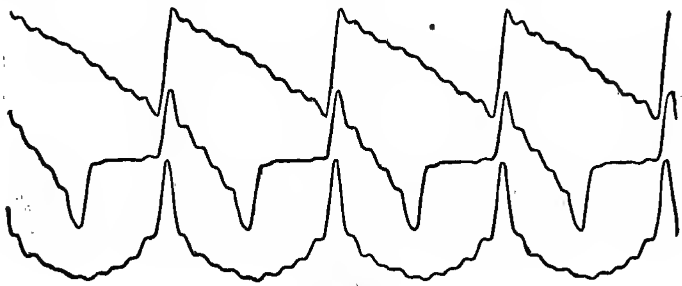

FIG. 91.

In Fig. 91 are shown three curves, each representing a fundamental accompanied by the harmonics up to the tenth. The

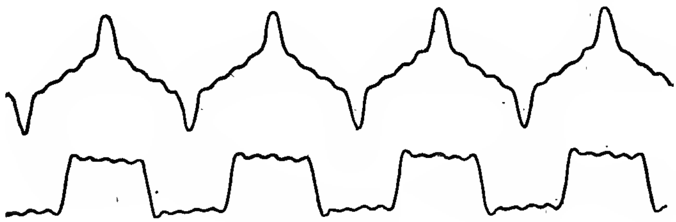

FIG. 92.

curves differ only in the different phases of the components relative to each other.

Fig. 92 shows similar curves produced by a fundamental accompanied by the odd harmonics.

263. Resonators for the Study of Complex Tones.—An apparatus devised by Helmholtz serves to analyze complex tones and indicate the simple tones of which they are composed. It consists of a series of hollow spheres or cylinders, called *resonators*, which are tuned to certain tones. If a tube lead from the resonator to the ear and a sound be produced, one of

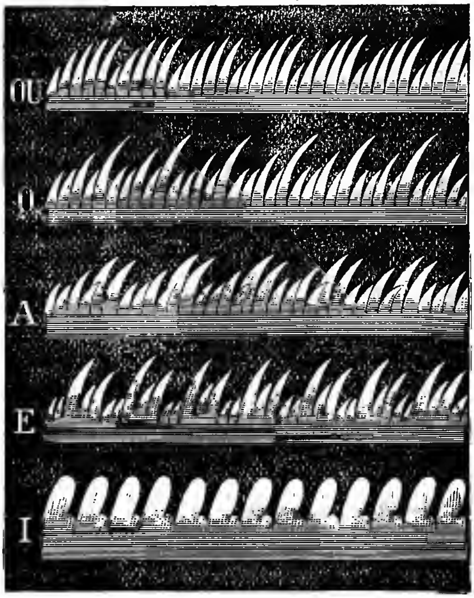

FIG. 93.

the components of which is the tone to which the resonator is tuned, the mass of air in it will be set in vibration and that tone will be clearly heard; or, if the resonator be connected by a rubber tube to a manometric capsule (§ 241), the gas flame connected with the capsule will be disturbed whenever the tone to which the resonator is tuned is produced in the vicinity, either by itself or as a component of a complex tone. By trying the resonators of a series, one after another, the several components of a complex tone may be detected and its composition demonstrated.

264. Vowel Sounds.—Helmholtz has shown that the differences between the vowel sounds are only differences of quality. That the vowel sounds correspond to distinct forms of vibration is well shown by means of the manometric flame. By connecting a mouthpiece to the rear of the capsule, and singing into it the different vowel sounds, the flame images assume distinct forms for each. Some of these forms are shown in Fig. 93.

265. Optical Method of Studying Vibrations. — The vibratory motion of sounding bodies may sometimes be studied

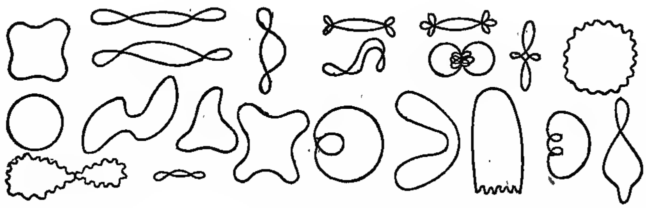

Fig. 94.

to advantage by observing the lines traced by luminous points upon the vibrating body or by observing the movement of a beam of light reflected from a mirror attached to the body.

Young studied the vibrations of strings by placing the string where a thin sheet of light would fall across it, so as to illuminate a single point. When the string was caused to vibrate, the path of the point appeared as a continuous line, in consequence of the persistence of vision. Some of the results which he obtained are given in Fig. 94, taken from Tyndall on Sound.

The most interesting application of this method was made by Lissajou to illustrate the composition of vibratory motions at right angles to each other. If a beam of light be reflected to a screen from a mirror attached to a tuning-fork, when the tuning-fork vibrates the spot on the screen will describe a simple harmonic motion and will appear as a straight line of light. If the beam, instead of being reflected to a screen, fall upon a mirror attached to a second fork, mounted so as to vibrate in

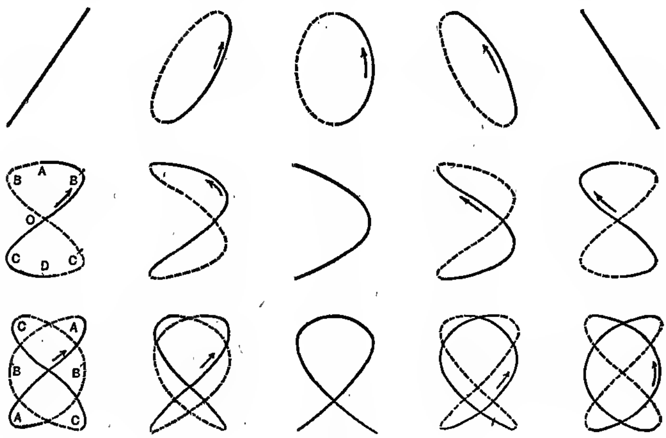

FIG. 95.

a plane at right angles to the first, the spot of light will, when both forks vibrate, be actuated by two simple harmonic motions at right angles to each other and the resultant path will appear as a curve more or less complicated, depending upon the relation of the two forks to each other as to both period and phase (§ 19). Fig. 95 shows some of the simpler forms of these curves. The figures of the upper line are those produced by two forks in unison; those of the second line by two forks of which the vibration numbers are as 2 : 1; those of the lower line by two forks of which the vibration numbers are as 3 : 2.

CHAPTER V.

EFFECTS OF THE COEXISTENCE OF SOUNDS.

266. Beats.—It has already been explained (§ 250) that, when two tones of nearly the same pitch are sounded together, variations of intensity, called *beats*, are heard. Helmholtz's theory of the perception of beats was, that, of the little fibres in the ear which are tuned so as to vibrate with the various tones, those which are nearly in unison affect one another so as to increase and diminish one another's motions, and hence that no beats could be perceived unless the tones were nearly in unison. Beats are, however, heard when a tone and its octave are not quite in tune, and, in general, a tone making n vibrations produces m beats when sounded with a tone making $2n \pm m$, $3n \pm m$, etc., vibrations. This was explained in accordance with Helmholtz's theory, by assuming that one of the harmonics of the lower tone, which is nearly in unison with

I

II

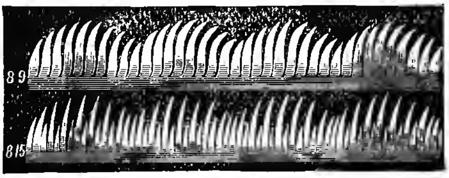

FIG. 96.

the upper, causes the beats, or, in cases where this is inadmissible, that they are caused by the lower tone in conjunction with a resultant tone (§ 267). An exhaustive research by König, however, has demonstrated that beats are perceived when neither of the above suppositions is admissible. Figs. 96 and 97 show that the resultant vibrations are affected by changes of amplitude similar to, though less in extent than, the changes which occur when the tones are nearly in unison. In Fig. 96, I represents a flame image obtained when two tones making n and $n \pm m$ vibrations respectively, are produced to-

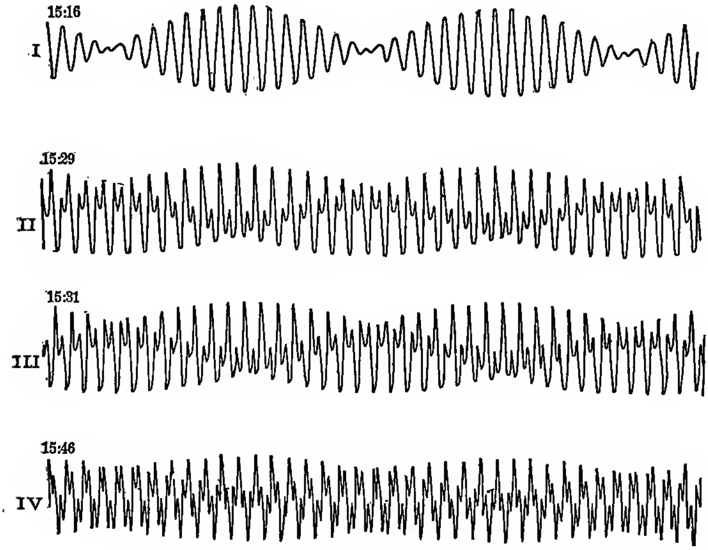

FIG. 97.

gether, and II represents the image when the number of vibrations are n and $2n \pm m$. Fig. 97 shows traces obtained mechanically. In I the numbers of the component vibrations were n and $n + m$, in II and III n and $2n \pm m$, and in IV n and $3n + m$. In all these cases a variation of amplitude occurs during the same intervals, and it seems reasonable to suppose that those variations of amplitude should cause variations in intensity in the sound perceived.

Cross has shown that the beating of two tones is perfectly well perceived when the tones themselves are heard separately by the two ears; one tone being heard directly by one ear, while the other, produced in a distant room, is heard by the other ear by means of a telephone. Beats are also perceived when tones are produced at a distance from each other and from the listener, who hears them by means of separate telephones through separate lines. In this case there is no possibility of the formation of a resultant wave, or of any combination of the two sounds in the ear.

267. Resultant Tones.—*Resultant tones* are produced by combinations of two tones. Those most generally recognized have a vibration number equal to the sum or difference of the vibration numbers of their primaries. For instance, ut_6, making 2048 vibrations, and re_6, making 2304 vibrations, when sounded together give ut_3, making 256 vibrations. These tones are only heard well when the primaries are loud, and it requires an effort of the attention and some experience to hear them at all. Summation tones are more difficult to recognize than difference tones, nevertheless they have an influence in determining the general effect produced when musical tones are sounded together. Other resultant tones may be heard under favorable conditions. As described above, two tones making n and $n + m$ vibrations respectively, when m is considerably less than n, give a resultant tone making m vibrations, but a tone making n vibrations in combination with one making $2n + m$, $3n + m$, or $xn + m$ vibrations, gives the same resultant. This has sometimes been explained by assuming that intermediate resultants are produced, which, with one of the primaries, produce resultants of a higher order. In the case of the two tones making n and $3n + m$ vibrations, for instance, the first difference tone would make $2n + m$ vibrations. This tone and the one making n vibrations would give the tone making $n + m$ vibrations; this tone, in turn, and the one making n vibrations would give the tone making m vibrations. This last tone is the one which is heard most plainly, and it seems difficult to admit that it can be the resultant of tones which are only heard very feebly, and often not at all. In Fig. 97 are represented the resultant curves produced in several of these cases. The first curve corresponds to two tones of which the vibration numbers are as 15 : 16. It shows the periodic increase and decrease in amplitude, occurring once every 15 vibrations, which, if not too frequent, give rise to beats (§ 250). If the pitch of the primaries be raised, preserving the relation 15 : 16, the beats become more frequent, and finally a distinct tone is heard, the vibration number of which corresponds to the number of beats that should exist. It was for a long time considered that the resultant tone was merely the rapid recurrence of beats. Helmholtz has shown by a mathematical investigation that a distinct wave making m vibrations will result from the coexistence of two waves making n and $n + m$ vibrations, and he believes that mere alternations of intensity, such as constitute beats, occurring ever so rapidly cannot produce a tone.

In II and III (Fig. 97) are the curves resulting from two tones, the intervals between which are respectively

$$15 : 29 (= 2 \times 15 - 1) \quad \text{and} \quad 15 : 31 (= 2 \times 15 + 1).$$

Running through these may be seen a periodic change corresponding exactly in period to that shown in I. The same is true also of the curve in IV, which is the resultant for two tones the interval between which is $15 : 46 (= 3 \times 15 + 1)$. In all these cases, as has been already said (§ 266), if the pitch of the components be not too high, one beat is heard for every 15 vibrations of the lower component. Fig. 96 shows the flame images for the intervals $n : n + m$ and $n : 2n + m$. The varying amplitudes resulting in m beats per second are very evident in both. In all these cases, also, as the pitch of the components rises the beats become more frequent, and finally a resultant tone is heard, having, as already stated, one vibration for every 15 vibrations of the lower component. In Fig. 98

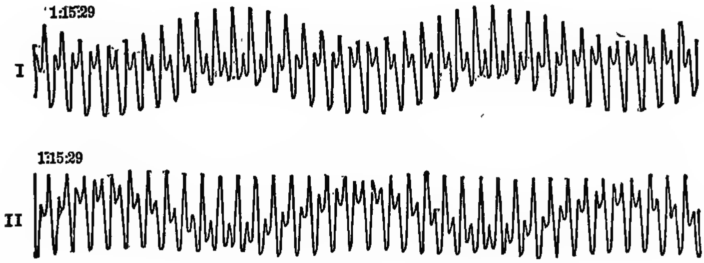

FIG. 98.

are shown two resultant curves having three components of which the vibration numbers are as 1 : 15 : 29. In I the three components all start in the same phase. In II, when 15 and 29 are in the same phase, I is in the opposite phase.

CHAPTER VI.

VELOCITY OF SOUND.

268. Theoretical Velocity.—The disturbance of the parts of any elastic medium which is propagating sound is assumed, in theoretical discussions, to take place in the line of direction of the propagation of the sound, and to be such that the type of the disturbance remains unaltered during its propagation. The velocity of propagation of such a disturbance may be investigated by the following method, due to Rankine.

Let us consider, as in § 242, a portion of the elastic medium in the form of an indefinitely long cylinder. If a disturbance be set up at any cross-section of this cylinder (Fig. 99), which consists of a displacement of the matter in that cross-section in the direction of the axis of the cylinder, it will, by hypothesis, be propagated in the direction of the axis with a constant velocity V, which is to be determined. If we consider any cross-section of the cylinder which is traversed by the disturbance, the matter which passes through it at any instant will

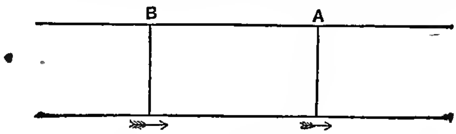

FIG. 99.

have a velocity which may vary from zero to the maximum velocity of the vibrating matter, either positively when this velocity is in the direction of propagation of the disturbance, or negatively when it is opposite to it.

If we now conceive an imaginary cross-section A to move along the cylinder with the disturbance with the velocity V, the velocity of the particles in it at any instant will be always the same. Let us call this velocity v_a. The velocity of the cross-section relative to the moving particles in it is then $V - v_a$. If we represent by d_a the density of the medium at the cross-section through which the velocity of the particles is v_a, which is the same for all positions of the moving cross-section, and if we assume that the area of the cross-section is unity, then the quantity of matter M which passes through the moving cross-section in unit time is

$$M = d_a(V - v_a).$$

If we conceive any other cross-section B to be moving with the disturbance in a similar manner, the same quantity of matter M will pass through it in unit time, since the two cross-sections move with the same velocity and the density of the matter between them remains the same. Hence we have $M = d_b\,(V - v_b)$, where d_b and v_b represent the quantities at the cross-section B corresponding to those at the cross-section A represented by d_a and v_a. Hence $d_a(V - v_a) = d_b(V - v_b)$. Since this equation is true whatever be the distance between the cross-sections, it is true for that position of the cross-section B for which $v_b = 0$, and for which $d_b = D$, the density of the medium in its undisturbed condition. Hence we have $M = DV$, $d_a(V - v_a) = DV$, and

$$\frac{v_a}{V} = \frac{d_a - D}{d_a}. \tag{107}$$

If the disturbance be small, the expression on the right is approximately the condensation per unit volume of the medium at the cross-section A, and the equation shows that the ratio of the velocity of the matter passing through the cross-section A to the velocity of propagation of the disturbance is equal to the condensation at that cross-section.

Now, to eleminate the unknown expressions v_a and d_a, we must find a new equation involving them. A quantity of matter M enters the region between the two moving cross-sections with the velocity v_a, and an equal quantity leaves the region with the velocity v_b. The difference of the momenta of the entering and outgoing quantities is $M(v_a - v_b)$. This difference can only be due to the different pressures p_a and p_b on the moving cross-sections, since the interactions of the portion of matter between those cross-sections cannot change the momentum of that portion. Hence we have

$$M(v_a - v_b) = p_a - p_b.$$

If we for convenience assume $v_b = 0$, we have $p_b = P$, the pressure in the medium in its undisturbed condition. If we further substitute for v_a its value, we obtain $MV = d_a \frac{p_a - P}{d_a - D}$. If the changes in pressure and density be small, the quantity $d_a \frac{p_a - P}{d_a - D}$ equals E, the modulus of elasticity of the medium. If we further substitute for M its value VD, we obtain finally

$$V^2 = \frac{E}{D} \text{ or } V = \sqrt{\frac{E}{D}}. \qquad (108)$$

269. Velocity of Sound in Air.—In air at constant temperature the elasticity is numerically equal to the pressure (§ 77). The compressions and rarefactions in a sound-wave occur so rapidly that during the passage of a wave there is no time for the transfer of heat, and the elasticity to be considered, therefore, is the elasticity when no heat enters or escapes (§ 158).

If the ratio of the two elasticities be represented by γ we have for the elasticity when no heat enters or escapes $E = \gamma P$, and the velocity of a sound-wave in air at zero temperature is given by

$$V = \sqrt{\frac{\gamma P}{D}}.$$

The coefficient γ equals 1.41. P is the pressure exerted by a column of mercury 76 centimetres high and with a cross-section of one square centimetre, or $76 \times 13.59 \times 981 = 1013373$ dynes per square centimetre. D equals 0.001293 grams at 0°, hence

$$V = \sqrt{\frac{1.41 \times 1013373}{0.001293}} = 33240,$$

or 332.4 metres per second.

Since the density of air changes with the temperature, the velocity of sound must also change. If d_t represent the density at temperature t, and d_0 the density at zero,

$$d_t = \frac{d_0}{1 + kt},$$

from § 128. The formula for velocity then becomes

$$V = \sqrt{\frac{\gamma P}{d_0}(1 + kt)}.$$

This formula shows that the velocity at any temperature is the velocity at 0° multiplied by the square root of the factor of expansion.

270. Measurements of the Velocity of Sound.—The velocity of sound in air has been measured by observing the time required for the report of a gun to travel to a known distance.

One of the best determinations was that made in Holland in 1822. Guns were fired alternately at two stations about nine miles apart. Observers at one station observed the time of seeing the flash and hearing the report from the other. The guns being fired alternately, and the sound travelling in opposite directions, the effect of wind was eliminated in the mean of the results at the two stations. It is possible, by causing the sound-wave to act upon diaphragms, to make it record its own time of departure and arrival, and by making use of some of the methods of estimating very small intervals of time the velocity of sound may be measured by experiments conducted within the limits of an ordinary building.

The velocity of sound in water was determined on Lake Geneva in 1826 by an experiment analogous to that by which the velocity in air was determined.

In § 255 and § 257 it is shown that the time of one vibration of any body vibrating longitudinally is the time required for a sound-wave to travel twice the distance between two nodes. The velocity may, therefore, be measured by determining the number of vibrations per second of the sound emitted, and measuring the distance between the nodes.

In an open organ-pipe, or a rod free at both ends, when the fundamental tone is sounded the sound travels twice the length of the rod or pipe during the time of one complete vibration. If rods of different materials be cut to such lengths that they all give the same fundamental tone when vibrating longitudinally, the ratio of their lengths will be that of the velocity of sound in them.

In Kundt's experiment, the end of a rod having a light disk attached is inserted in a glass tube containing a light powder strewn over its inner surface. When the rod is made to vibrate longitudinally, the air-column in the tube, if of the proper length, is made to vibrate in unison with it. This agitates the powder and causes it to indicate the positions of the nodes in the vibrating air-column. The ratio of the velocity of sound in the solid to that in air is thus the ratio of the length of the rod to the distance between the nodes in the air-column.

LIGHT.

CHAPTER I.

PROPAGATION OF LIGHT.

271. Vision and Light.—The ancient philosophers, before Aristotle, believed that vision consisted in the contact of some subtle emanation from the eye with the object seen. Aristotle showed the absurdity of this view by suggesting that if it were true, one should be able to see in the dark. Since his time, it has been generally admitted that vision results from something proceeding from the body seen to the eye, and there impressing the optic nerve. This we call *light*.

Optics treats of the phenomena of light. It is conveniently divided into two branches, *Physical Optics*, which treats of the phenomena resulting from the propagation of light through space and through different media, and *Physiological Optics*, which treats of the sense of vision.

272. Theories of Light.—At the time of Newton, light was generally considered to consist of particles which were not those of ordinary matter, projected from a luminous body, and exciting vision by their impact on the retina. This theory was strongly supported by Newton himself, who found in it plausible explanations of most luminous phenomena then known. But even in Newton's time phenomena were known which could only be explained by assigning to the luminiferous particles very improbable forms and motions, and, since his time, facts have been discovered that are inconsistent with any emission theory.

The *undulatory* theory, which is the one universally adopted, assumes that light is a wave motion in an elastic medium pervading all space. All luminous bodies excite in this medium systems of waves which are propagated according to the same mechanical laws as those which govern wave systems in other media, some of which have been developed in § 19 and §§ 242–245. The undulatory theory has stood well the test of explaining newly discovered phenomena, and has moreover led to the discovery of phenomena not before known. The objections to the theory are that it requires the hypothesis of a medium of the existence of which there is no direct evidence, pervading space, and requires us to ascribe to that medium properties unlike those of any body with which we are acquainted.

A modified form of the undulatory theory, known as the *electromagnetic* theory of light, was proposed by Maxwell. It will be briefly presented after the facts connecting light and electricity have been considered.

273. Wave Surfaces.—In § 243 is explained the general mode of propagation of wave motion in accordance with Huyghens' principle. When light, emanating from a point, proceeds with the same velocity in all directions, the wave fronts are evidently concentric spherical surfaces. There are, however, many cases, especially in crystalline bodies, of unequal velocities in different directions. In these cases the wave fronts are not spherical but ellipsoidal, or surfaces of still greater complexity.

274. Straight Lines of Light.—When a small screen A (Fig. 100) is placed between the eye and a luminous point, the luminous point is no longer visible. Light cannot reach the eye by the curved or broken line PAE, and is therefore said to move in straight lines. This seems not to accord with Huyghens' principle which makes any wave front the resultant of an infinite number of elementary waves proceeding from the

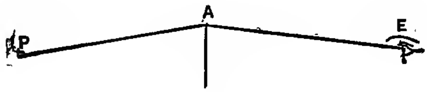

FIG. 100.

various points of the same wave front in one of its earlier positions. It can, however, easily be shown that when the wave lengths are small, the disturbance at any point P (Fig. 101) is due almost wholly to a very small portion of the approaching wave. Let us consider first the case of an isotropic medium, in which light moves in all directions with equal velocities. Let mn be the front of a plane wave perpendicular to the plane of the paper, moving from left to right or towards P. Draw PA perpendicular to the wave front, and draw Pa, Pb, etc., at such obliquities that Pa shall exceed PA by half a wave length, Pb exceed Pa by half a wave length, etc. We will designate the wave length by λ.

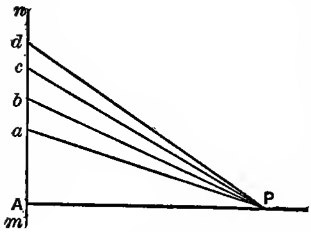

FIG. 101.

It is evident that the total effect at P will be the sum of the effects due to the small portions Aa, ab, etc. Since Pa is half a wave length greater than PA, and Pb half a wave length greater than Pa, each point of ab is half a wave length farther from P than some point in Aa; hence elementary waves from ab will meet at P waves from Aa in the opposite phase. It appears, therefore, that the effects at P of the portions ab and Aa are opposite in sign, and tend to annul each other. The same is true of bc and cd. But the effects of Aa and ab may be considered as proportional to their lengths. Hence, by computing the lengths, we can determine the resultant effect at P. Let $AP = x$. From the construction, we have

$$
\begin{aligned}
Aa &= \sqrt{\left(x+\frac{\lambda}{2}\right)^2 - x^2} = \sqrt{x\lambda + \frac{\lambda^2}{4}};\\
Ab &= \sqrt{(x+\lambda)^2 - x^2} = \sqrt{2x\lambda + \lambda^2};\\
Ac &= \sqrt{(x+\tfrac{3}{2}\lambda)^2 - x^2} = \sqrt{3x\lambda + \tfrac{9}{4}\lambda^2};\\
Ad &= \sqrt{(x+2\lambda)^2 - x^2} = \sqrt{4x\lambda + 4\lambda^2};\\
\text{etc.} &= \quad \text{etc.}
\end{aligned}
$$

For light the values of λ are between 0.00039 and 0.00076 mm., and if x be taken as 1000 mm., λ^2 will be very small in comparison to $x\lambda$ and may be omitted. The above formulas then become, if $\sqrt{x\lambda}$ be represented by l,

$$
\begin{aligned}
Aa &= l\sqrt{1};\\
Ab &= l\sqrt{2};\\
Ac &= l\sqrt{3};\\
Ad &= l\sqrt{4};\\
\text{etc.} &= \text{etc.},
\end{aligned}
$$

and the several portions into which the wave front is divided are

$$
\begin{aligned}
Aa &= l &&= 1l;\\
ab &= l(\sqrt{2} - 1) &&= 0.414l;\\
bc &= l(\sqrt{3} - \sqrt{2}) &&= 0.318l;\\
cd &= l(\sqrt{4} - \sqrt{3}) &&= 0.268l.
\end{aligned}
$$

Taking now the pairs of which the effects at P are opposite in sign, we find Aa a little more than twice ab, while bc and cd are nearly equal. It is evident also, that for portions beyond d, adjacent pairs will be still more nearly equal, and the effect at P, therefore, of each pair of segments beyond b almost vanishes. The effect at P is then almost wholly due to that portion of Aa that is not neutralized by ab. But, taking the greatest value of λ, $Aa = \sqrt{x\lambda} = \sqrt{0.76} = 0.87$ mm., a very small distance. Hence, under the conditions assumed, the effect at any point P is due to that portion of the wave front near the foot of the perpendicular let fall from P on the wave front. It may be demonstrated by experiment that the portions of the wave beyond Aa neutralize each other. Suppose a screen mn in the position shown in Fig. 102. The point P will be in shadow. If the darkness at P is due to interference as explained, light should be restored by suppressing the interfering waves. If a second screen be placed at $m'n'$ so as to cut off the waves proceeding from points above b, waves from points between a and b will no longer be neutralized, and light should fall at P. To test this conclusion the edge of a flat flame may be observed through a narrow slit in a screen. Instead of the narrow edge of the flame, a broad luminous surface is seen, in which the brightness gradually diminishes from the centre towards the edges. If we consider the wave front just entering the slit, it will be seen that elementary waves proceed from all points of it, and the slit being very narrow it is only in very oblique directions that pairs of these waves can meet in opposite phases. Hence, light proceeds in oblique lines behind the screen, and from our habit of locating visible objects back along the line of light entering the eye, the flame appears as a

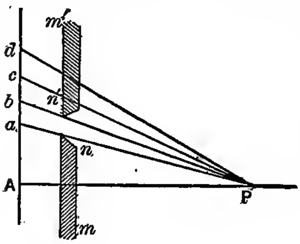

FIG. 102.

broad surface. It will be seen by reference to Fig. 101 that the elementary wave that first reaches P is the one to which the disturbance there is principally due. Other waves arriving later find there the opposite phase of some wave that has preceded them. When the velocity in all directions is the same, the first wave to reach P is the one that starts from the foot of a perpendicular let fall from P on the wave front. Hence light is said to travel in straight lines perpendicular to the wave front. If, however, light does not move with equal velocities in all directions, the last statement is no longer true, as will be seen from Fig. 103. Here mn represents a wave front, proceeding towards P in a medium in which the velocities in different directions are such that the elementary wave surfaces are ellipsoids. The ellipses in the figure may be taken as sections of these ellipsoids. The wave first to reach P is not the one that starts from A at the foot of the perpendicular, but from A'. It is from A' that P derives its light, and the line of propagation is no longer perpendicular to the wave front.

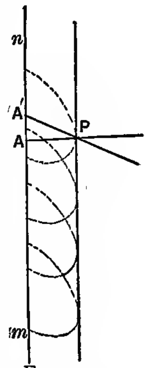

FIG. 103.

It is important to note that the deductions of this section apply only where λ is small in relation to x, so that λ^2 may be neglected in comparison with $x\lambda$. With sound-waves this is not true, and if a computation similar to that given above for light-waves be made for sound, not omitting λ^2, it will be seen why there are no definite straight lines of sound and no sharp acoustic shadows.

275. Principle of Least Time.—The above are only particular cases of a law of very general application, that light in going from one point to another follows the path that requires least time. The reason is that values in the vicinity of a minimum change slowly, and there will be a number of points in the neighborhood of that point from which the light-waves are

26

propagated to the given point in the least time, from which waves will proceed to that point in sensibly the same time, and, meeting in the same phase, combine to produce light. It is also true that values change slowly in the vicinity of a maximum, and there are cases where the path followed by the light is determined by the fact that the time is a maximum instead of a minimum.

276. Shadows.—An optical *shadow* is the space from which light is excluded by an opaque body. When the luminous source is a point, or very small, the boundary between the light and shadow is very sharp. When the luminous source is large, there is a portion of the space behind the opaque body, called the *umbra*, which is in deep shadow, and surrounding this is a space which is in shadow with reference to one portion of the luminous source while it is in the light with reference to another portion. The space from which light is only partially excluded is the *penumbra*. Fig. 104 shows the boundaries of the umbra and penumbra. It is evident that the light diminishes gradually from the outer boundary of the penumbra to the boundary of the umbra.

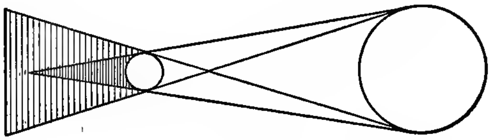

FIG. 104.

277. Images by Small Apertures.—If light from a single luminous point pass through a small hole of any form, and fall on a screen at some distance, it produces a luminous spot of the same form as the opening. Light from several points will produce several such spots. If the luminous source be a surface, the spots produced by the light from its several points will overlap each other and form an illuminated surface, which, if the source be large in comparison with the opening, will have the general form of the source, and will be inverted. The illuminated surface is an inverted *image* of the source. If a small opening be made in the window-shutter of a darkened room, images of external objects will be seen on the wall opposite. The smaller the opening, the more sharply defined, but the less brilliant, is the image.

CHAPTER II.

REFLECTION AND REFRACTION.

278. Law of Reflection.—In § 246, it is shown that when a wave passes from one medium into another where the particles constituting the wave move with greater or less facility, a wave is propagated back into the first medium. It is shown in § 247, that when the surface separating the two media is a plane surface, the centres of the incident and reflected waves are on the same perpendicular to the surface, and at equal distances on opposite sides. Considering the lines to which, as shown in § 274, the wave propagation in the case of light is restricted, a very simple law follows at once from this relation of the incident and reflected waves. In

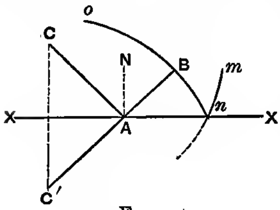

FIG. 105.

Fig. 105, C and C' represent the centres of the incident and reflected waves mn, on. CA, AB are the paths of the incident and reflected light. It will be evident from the figure that CA, AB are in the same plane normal to the reflecting surface, and that they make equal angles with the normal AN. CAN is called the angle of incidence, and NAB the angle of reflection. Hence we may state the *law of reflection* as follows: The angles of incidence and reflection are equal, and lie in the same plane normal to the reflecting surface. It can easily be shown that light traverses the path CAB from C to B which fulfils these laws, in less time than it requires to traverse any other path by way of the reflecting surface.

279. Law of Refraction.—If the incident wave pass from the one medium into the other, there is in general a change in the wave front, and a consequent change in the direction of the light. Let us first consider the simple case of a plane wave entering a homogeneous, isotropic medium of which the bounding surface is plane. Suppose both planes perpendicular to the

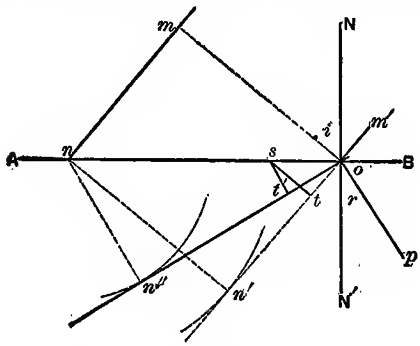

FIG. 106.

plane of the paper, and let AB (Fig. 106) represent the intersection of the surface of the medium, and mn the insection of the wave, with that plane. Let v represent the velocity of light in the medium above AB, and v' the velocity in the medium below it. Let $m'o$ be the position of the wave in the first medium after a time t. Then mo equals vt. As the wave front passes from mn to $m'o$, the points of the separating surface between n and o are successively disturbed, and become centres of spherical waves propagated into the second medium with the velocity v'. The wave surface of which the centre is n would, at the end of time t, have a radius $nn'' = v't$, such that $\frac{nn'}{nn''} = \frac{v}{v'}$. Similarly, the wave from any other point, as s, would have a radius st' such that $\frac{st}{st'} = \frac{v}{v'}$, and the wave surface within the second medium is evidently the plane on''. As the direction of propagation is perpendicular to the wave front, op will represent the direction of the light in the second medium. In the triangles non' and non'' we have $nn' = no \sin Aon$, and $nn'' = no \sin Aon''$; hence

$$\frac{\sin Aon'}{\sin Aon''} = \frac{nn'}{nn''} = \frac{v}{v'}.$$

If we represent the *angle of incidence moN* by i, and the *angle of refraction poN'* by r, we have

$$\frac{\sin i}{\sin r} = \frac{v}{v'}, = \mu, \text{ a constant.} \qquad (109)$$

This constant is called the *index of refraction.* This is the expression of Snell's *law of refraction.* Here again the time required for the light to pass by mop from m in one medium to p in the other is less than by any other path.

We may now trace a wave through a medium bounded by plane surfaces. Suppose the wave front and bounding planes of the medium all perpendicular to the plane of the paper. We shall have as above for the first surface $\frac{\sin i}{\sin r} = \frac{v}{v'}, = \mu$, and for the second surface $\frac{\sin i'}{\sin r'} = \frac{v'}{v''} = \mu'$.

If, as is often the case, the light emerge into the first medium,

$$v'' = v, \quad \text{and} \quad \frac{\sin i'}{\sin r'} = \frac{v'}{v} = \frac{1}{\mu}. \qquad (110)$$

If the bounding planes be parallel, $i' = r$, and we have

$$\frac{\sin r}{\sin r'} = \frac{1}{\mu};$$

hence $i = r'$, or the incident and emergent waves are parallel. If the two bounding planes form an angle A the body is called a *prism*. The wave incident upon the second face will make with it an angle $A - r$, and the emergent wave is found by the relation

$$\frac{\sin(A-r)}{\sin r'} = \frac{1}{\mu} \text{ or } \frac{\sin r'}{\sin(A-r)} = \mu.$$

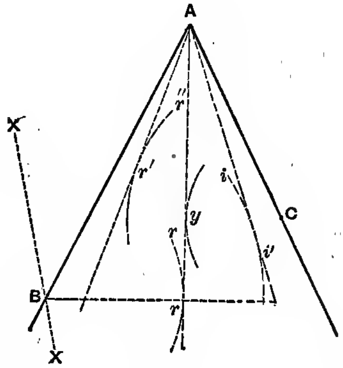

FIG. 107.

The direction of the emerging wave front may be found by construction.

Draw Ai (Fig. 107) parallel to the incident wave. From some point B on AB describe an arc tangent to Ai; from the same point with a radius $\frac{Bi}{\mu}$ describe the arc rr. Ar, tangent to rr, is the refracted wave front. From some point C on AC describe an arc tangent to Ar, and from the same point as centre describe another arc $r'r'$ with a radius $\mu \times Cy$. A tangent from A to $r'r'$ is parallel to the emergent wave. It might be that A would fall inside the arc $r'r'$ so that no tangent could be drawn. That would mean that there could be no emergent wave. The angle of incidence for which this occurs can readily be obtained from Eq. (110). We have

$$\frac{\sin i'}{\sin r'} = \frac{1}{\mu}, \quad \text{or} \quad \sin r' = \mu \sin i'.$$

Now the maximum value of $\sin r'$ is 1, which is reached when $\sin i' = \frac{1}{\mu}$. Any larger value of $\sin i'$ gives an impossible value for sin r'. The angle $i = \sin^{-1}\frac{1}{\mu}$ is called the *critical angle* of the substance. For larger angles of incidence the light cannot emerge, but is *totally reflected* within the medium.

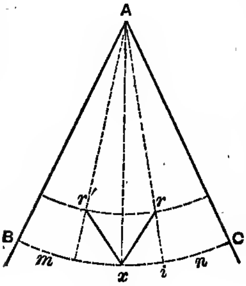

FIG. 108.

Another construction for the front of the emergent wave is very instructive. Let AB, AC (Fig. 108), be the faces of the prism, and let Ai drawn through A be parallel to the front of the incident wave. With A as centre, and any radius, draw an arc im. From the same centre with radius $Ar = \frac{Ai}{\mu}$ describe another arc. From r draw rx parallel to AB and join Ax. Ax is parallel to the front of the refracted wave. For in the triangle Arx we have

$$\frac{\sin Arx}{\sin Axr} = \frac{\sin irx}{\sin Axr} = \frac{Ax}{Ar} = \mu, \text{ by construction.}$$

Since irx equals the angle of incidence, Axr equals the angle of refraction. Now draw xr' parallel to AC, and Ar' is parallel to the front of the emergent wave. The angle $r'Ar$ is the deviation that the wave suffers in passing through the prism. Suppose the prism to rotate about A and the angle of incidence to change in such a way that the condition of things may be always represented by rotating the angle rxr', of which the sides are parallel to the sides of the prism, around x. It is plain that the arc $r'r$ will be longer or shorter as it crosses the angle more or less obliquely, and that its length will be a minimum when xr' and xr are equal—that is, when the line Ax bisects the angle at x and consequently the angle A of the prism. But the arc $r'r$ may be taken as the measure of the angle of deviation $r'Ar$ at its centre. Hence that angle is a minimum when it is bisected by Ax, and when, therefore, the angles of incidence and of emergence are equal. Considering that the path of the light is perpendicular to the wave front, the above construction shows that the deviation, when μ is greater than unity, is always toward the thicker portion of the prism. The case when emergence is no longer possible is also shown by the failure of xr', parallel to AC, to cut the arc $r'r$. The critical angle is reached when xr' becomes tangent to $r'r$. If, in a prism of any substance, xr and xr' be both tangent to $r'r$, the angle of that prism is the greatest angle which will admit of the passage of light through the prism.

If a beam of white light be allowed to fall upon a prism through a narrow slit, it will be refracted, in general, in accordance with the law already given. The image of the slit, however, when projected upon a screen, appears not as a single line of white light, but as a variously colored band. This is due to the fact that the indices of refraction for light of different colors are different. Hence the index of refraction of a substance, as ordinarily given, depends upon the color of the light used in determining it, and has no definite meaning unless that color is stated.

280. Plane Mirrors.—The wave *on*, represented in Fig. 105, is the same as would have come from a luminous point at C' if the reflecting surface did not intervene. If this wave reach the eye of an observer, it has the same effect as though coming from such a point, and the observer apparently sees a luminous point at C'. C' is a *virtual image* of C. When an object is in front of a plane mirror each of its points has an image symmetrically situated in relation to the mirror, and these constitute an image of the object like the latter in all respects, except that by reason of symmetry it is reversed in one direction.

The reflected light may for all purposes be considered as coming from the image. If it fall on a second mirror and be again reflected, a second image appears behind this mirror, the position of which is determined by considering the first image as an object. When two mirrors make an angle, an object between them will have a series of images, as shown in Fig. 109. *AB* and *AC* represent the intersections of the two mirrors with the plane of the paper, to which they are supposed perpendicular. *O* is the object. It will have an image produced by *AB*, the position of which is found by drawing *Ob* perpendicular to *AB* and making $mb = mO$. The light reflected from *AB* proceeds as though *b* were the object, and falling on *AC* is again reflected, giving an image at c'. Proceeding from *AC*, it may suffer a third reflection from *AB* and give a third image at b''.

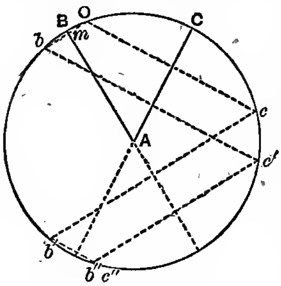

FIG. 109.

With the angle as in the figure none of the light can suffer a fourth reflection, because after the third reflection the light proceeds as though originating at b'', and b'' is behind the plane of the mirror *AC*. Images c, b', and c'' are produced by light which suffers its first reflection from *AC*. It is easy to show that all these points are equidistant from *A*, and hence are on the circumference of a circle of which *A* is the centre. If *OAC* were an even aliquot part of four right angles, c'' and b'' would coincide, and the whole number of images, including the object, would be the quotient of four right angles by the angle formed by the mirrors. This is the principle of the kaleidoscope.

281. Spherical Mirrors.—A spherical mirror is a portion of a spherical surface. It is a concave mirror if reflection occur on the concave or inner surface; a convex mirror if it occur on the convex surface. The centre of the sphere of which the mirror forms a part is its *centre of curvature*. The middle point of the surface of the mirror is the *vertex*. A line through the centre of curvature and the vertex is the *principal axis*. Any other line through the centre of curvature is a *secondary axis*. The angle between radii drawn to the edge of the mirror on opposite sides of the vertex is the *aperture*. To investigate the effects of reflection from a spherical surface, let us consider first a concave mirror. Let a light-wave emanate from a point L on the principal axis (Fig. 110). In general,

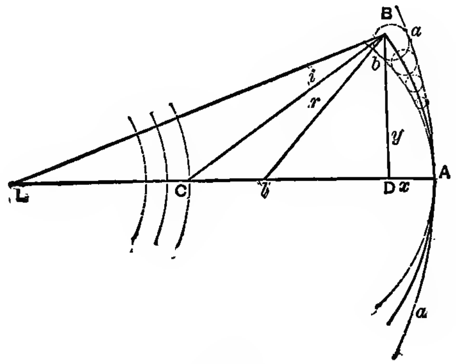

FIG. 110.

different points of the wave will reach the mirror successively, and, considering the elementary waves that proceed in turn from its several points, the reflected wave surface may be constructed as for a plane mirror. If the mirror were not there the wave front would, at a certain time, occupy the position *aa*. Drawing the elementary wave surfaces we have *bb*, the position at that instant of the reflected wave. Its form suggests that of a spherical surface, concave toward the front, and having a centre at some point l on the axis. If we assume it to be so, and try to determine by analysis the position of l, a real definite result will be proof of the correctness of our assumption. If *bb* be a spherical surface and l its centre, it is plain that the disturbances propagated from the various points of bb will reach l at the same instant, and l will at that instant be the wave front. It is plain, too, that the time occupied by the wave in going from the radiant point to all points of the same wave front must be the same. Hence, in a homogeneous medium, the length of path to the various points of the wave must be constant, that is, in the case under consideration, $LB + Bb$ must be constant for all points of the wave front bb. If l be a subsequent position of bb, it follows that $LB + Bl$ must be constant wherever the point B is situated on the reflecting surface. Draw BD perpendicular to the axis of the mirror. Represent BD by y, AD by x, LA by p, lA by p', and CA by r. Then we have $LB = \sqrt{(p-x)^2 + y^2}$, and $y^2 = (2r - x)x = 2rx - x^2$. Hence follows

$$\begin{aligned} LB &= \sqrt{p^2 - 2px + x^2 + 2rx - x^2} \\ &= \sqrt{p^2 + 2x(r-p)}. \end{aligned}$$

If the aperture be small, x will be small in comparison with the other quantities, and we may obtain the value of LB to a near approximation by extracting the root of the expression found above and omitting terms containing the second and higher powers of x. We obtain

$$LB = p + \frac{x}{p}(r-p) + \ldots.$$

In like manner we have

$$lB = p' + \frac{x}{p'}(r-p') + \ldots,$$

whence $$LB + lB = p + p' + \frac{x}{p}(r-p) + \frac{x}{p'}(r-p').$$

When B coincides with A, the above value becomes $p + p'$, and since upon our supposition all values of $LB + lB$ are equal, we must have

$$p + p' = p + p' + \frac{x}{p}(r - p) + \frac{x}{p'}(r - p'),$$

from which we obtain

$$\frac{r}{p} + \frac{r}{p'} = 2$$

and
$$p' = \frac{pr}{2p - r}.$$

As this is a definite value, it follows that, for the apertures for which the approximations by which the result was arrived at are admissible, the wave surface is practically spherical. Since the disturbances propagated from bb reach l simultaneously, their effects are added, and the disturbance at l is far greater than at any other point. The effect of the wave motion is concentrated at l, and this point is therefore called a *focus*. Since the light passes through l, it is a *real focus*. If l were the radiant point, it is clear that the reflected light would be concentrated at L. These two points are therefore called *conjugate foci*. If we divide both sides of the equation $\frac{r}{p} + \frac{r}{p'} = 2$ by r, we have

$$\frac{1}{p} + \frac{1}{p'} = \frac{2}{r}, \qquad (111)$$

which is the usual form of the equation used to express the relation between the distances from the mirror of the conjugate foci.

A discussion of this equation leads to some interesting results. Suppose $p = \infty$, then $p' = \frac{1}{2}r$; that is, when the radiant is at an infinite distance from the mirror, the focus is midway between the mirror and the centre. In this case the incident wave is normal to the principal axis, and the focus is called the *principal focus.* Suppose $p = r$; $p' = r$ also. When $p = \frac{1}{2}r$, $p' = \infty$. When $p < \frac{r}{2}$, $\frac{1}{p} > \frac{2}{r}$ and $\frac{1}{p'} = \frac{2}{r} - \frac{1}{p} =$ a negative quantity. To interpret this negative result it should be remembered that all the distances in the formulas were assumed positive when measured from the mirror toward the

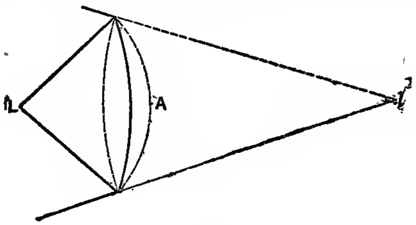

Fig. 111.

source of light. A negative result means that the distance must be measured in the opposite direction, or behind the mirror. Fig. 111 represents this case. It is evident that the reflected wave is convex toward the region it is approaching, and proceeds as though it had come from *l*. *l* is therefore a *virtual focus.* Either of the other quantities of the formula may have negative values. p will be negative if waves approaching their centre *l* fall on the mirror. Plainly they would be reflected to *L* at a distance from the mirror less than $\frac{r}{2}$, as may be seen from the formula. If r be negative, the centre is behind the mirror. The mirror is then convex, and the formula shows that for all positive values of p, p' is negative and numerically smaller than p.

282. Refraction at Spherical Surfaces.—The method of discussion which has been applied to reflection may be employed to study refraction at spherical surfaces. Let *BD* (Fig. 112) be a spherical surface separating two transparent

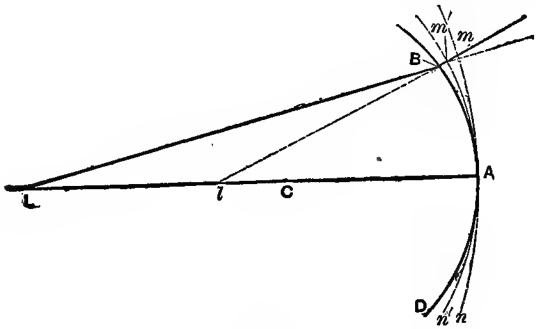

FIG. 112.

media. Let v represent the velocity of light in the first medium, to the left, and v' the velocity in the second medium, to the right, of *BD*. Let *L* be a radiant point, and *mn* a surface representing the position which the wave surface would have occupied at a given instant had there been no change in the medium, *m'n'* the wave surface as it exists at the same instant in the second medium in consequence of the different velocity of light in it. Assume as before, in § 281, that *m'n'* is a spherical surface with centre *l*. We have

$$\frac{Bm}{v} = \frac{Bm'}{v'},$$

and

$$\frac{LB + Bm}{v} = \frac{LB}{v} + \frac{Bm}{v} = C,$$

a constant for all points of *mn*. If *l* be the centre of the spherical surface *m'n'*, we have

$$\frac{lB}{v'} + \frac{Bm'}{v'} = C',$$

a constant for all points of $m'n'$.

Taking the difference of the last two equations, and remembering that

$$\frac{Bm}{v} = \frac{Bm'}{v'},$$

we obtain $$\frac{LB}{v} - \frac{lB}{v'} = C - C',$$

a constant for all points of BD, and hence

$$LB - \frac{v}{v'}lB = \text{a constant.}$$

But $\frac{v}{v'} = \mu$ is the index of refraction of the second substance in relation to the first. Hence $LB - \mu lB =$ a constant $= LA - \mu lA$. Using the notation of the last section, and substituting the values of LB and lB as there found, except that p'' is used instead of p', we have

$$p + \frac{x}{p}(r - p) - \mu\left(p'' + \frac{x}{p''}(r - p'')\right) = p - \mu p'',$$

whence we obtain $$\frac{r}{p} - \frac{\mu r}{p''} = 1 - \mu,$$

and $$\frac{\mu}{p''} - \frac{1}{p} = \frac{\mu - 1}{r}. \qquad (112)$$

If the medium to the right of BD be bounded by a second spherical surface, it constitutes a *lens*. Suppose this second surface to be concave toward l and to have its centre on AC. The wave $m'n'$, in passing out at this second surface, suffers a new change of form precisely analogous to that occurring at the first surface, and the new centre is given by the formula just deduced by substituting for p the distance of the wave centre from the new surface, and for μ the index of refraction of the third medium in relation to the second. If s represent the distance of l from the new surface, μ' the new index, and p' the new focal distance, we have

$$\frac{\mu'}{p'} - \frac{1}{s} = \frac{\mu' - 1}{r'}.$$

If we suppose the lens to be very thin we may put $s = p''$. If we suppose also that the medium to the right is the same as that to the left of the lens, μ' is equal to $\frac{1}{\mu}$. Hence

$$\frac{\frac{1}{\mu}}{p'} - \frac{1}{p''} = \frac{\frac{1}{\mu} - 1}{r'}.$$

Multiplying through by μ, we have

$$\frac{1}{p'} - \frac{\mu}{p''} = \frac{1-\mu}{r'} = -\frac{\mu - 1}{r'}.$$

Eliminating p'' between this equation and Eq. 112, we obtain

$$\frac{1}{p'} - \frac{1}{p} = (\mu - 1)\left(\frac{1}{r} - \frac{1}{r'}\right), \tag{113}$$

which expresses the relation between the conjugate foci of the lens. It should be noted that r in the above formulas represents the radius of the surface on which the light is incident, and r' that of the surface from which the light emerges. All the quantities are positive when measured toward the source of light. Fig. 113 shows sections of the different forms of

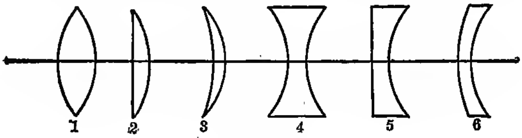

FIG. 113.

lenses produced by cominations of two spherical surfaces, or of one plane and one spherical surface.

An application of Eq. 113 will show that for the first three, which are thickest at the centre, light is concentrated, and for the second three diffused. The first three are therefore called *converging*, and the second three *diverging*, lenses. Let us consider the first and fourth forms as typical of the two classes. The first is a double convex lens. The r of Eq. 113 is negative because measured from the lens away from the source of light. The second term of the formula has therefore a negative value, and p' is negative except when $\frac{1}{p} > (\mu - 1)\left(\frac{1}{r} - \frac{1}{r'}\right)$. If $p = \infty$, we have $\frac{1}{p} = 0$ and $\frac{1}{p'} = (\mu - 1)\left(\frac{1}{r} - \frac{1}{r'}\right)$, a negative quantity because r is negative. p' is then the distance of the principal focus from the lens, and is called the *focal length* of the lens. The focal length is usually designated by the symbol f. Its negative value shows that the principal focus is on the side of the lens opposite the source of light. This focus is real, because the light passes through it. Eq. 113 is a little more simple in application if, instead of making the algebraic signs of the quantities depend on the direction of measurement, they are made to depend on the form of the surfaces and the character of the foci. If we assume that radii are positive when the surfaces are convex, and that focal distances are positive when foci are real, the signs of p' and r in Eq. 113 must be changed, since in the investigation p' is the distance of a virtual focus, and r the radius of a concave surface. The formula then becomes

$$\frac{1}{p'} + \frac{1}{p} = (\mu - 1)\left(\frac{1}{r} + \frac{1}{r'}\right). \qquad (114)$$

To apply this formula to a double concave lens, r and r' are both negative; p' is then negative for all positive values of p. That is, concave lenses have only virtual foci. For a plano-convex lens (Fig. 113, 2), if light be incident on the plane surface,

$$r = \infty \quad \text{and} \quad \frac{1}{p'} = (\mu - 1)\frac{1}{r} - \frac{1}{p}.$$

This gives positive values of p' and real foci for all values of

$$\frac{1}{p} < (\mu - 1)\frac{1}{r'}.$$

For a concavo-convex lens (Fig. 113, 6) the second member of the equation will be negative, since the radius of the concave surface is negative and less numerically than that of the convex surface. Hence p' is always negative and the focus virtual when L is real.

283. Images formed by Mirrors.—In Fig. 114 let ab represent an object in front of the concave mirror mn. We know from what precedes that if we consider only the light incident not too far from c, the light reflected will be concentrated at some point a' on the axis ac at a distance from the mirror given by Eq. 114. a' is a *real image* of a. In the same way b' is an image of b. If axes were drawn through other points of the object, the images of those points would be found in the same way. They would lie between a' and b', and $a'b'$ is therefore a real image of the object. It is inverted, and lies between the axes ac, bd, drawn through the extreme points of the object. The ratio of its size to that of the object is seen from the similar triangles abC, $a'b'C$, to be the ratio of the distances from C. From Eq. 111 we obtain

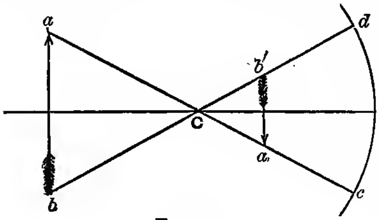

FIG. 114.

$$\frac{p'}{p} = \frac{r}{2p - r} = \frac{r - p'}{p - r}.$$

Since $r - p'$ and $p - r$ are respectively the distances from the centre of the image and object, we have

$$\frac{a'b'}{ab} = \frac{r - p'}{p - r} = \frac{p'}{p};$$

or, the image and object are to each other in the ratio of their respective distances from the mirror. As the object approaches, the image recedes from the mirror and increases in size. At the centre of curvature the image and object are equal, and when the object is within the centre and beyond the principal focus the image is outside the centre and larger than the object. When the object is between the principal focus and the mirror, the image is virtual and larger than the object. Convex mirrors produce only virtual images, which are erect and smaller than the object.

284. Images formed by Lenses.—Let us suppose an object in front of a double convex lens, which may be taken as a type of the converging lenses. The point c (Fig. 115) will have an image at the conjugate focus on the principal axis. a and b will have images on secondary axes drawn through those points respectively, and a point called the optical centre of the lens.

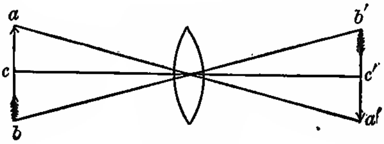

Fig. 115.

So long as these secondary axes make but a small angle with the principal axis, definite foci will be formed at the same distances as on the principal axis, and an image $a'b'$ will be formed which will be real and inverted, or virtual and erect, according to the distance of the object from the lens. The formula

$$\frac{1}{p}+\frac{1}{p'}=(\mu-1)\left(\frac{1}{r}+\frac{1}{r'}\right)=\frac{1}{f},$$

shows that when p increases p' diminishes, and conversely. It shows also that when p is less than f, p' is negative, and the image virtual. It is plain from the figure that the sizes of image and object are in the ratio of their distances from the lens. Diverging lenses, like diverging mirrors, produce only virtual images smaller than the object.

285. Optical Centre.—It was stated in the last section that the secondary axes of a lens pass through a point called the *optical centre*. The location of this point is determined as follows: In Fig. 116, let C, C' be the centres of curvature of the two surfaces of the lens, and let CA and $C'B$ be two parallel radii.

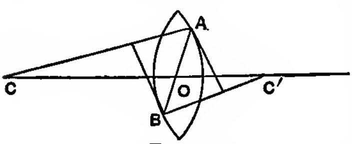

Fig. 116.

The tangents at A and B are also parallel, and light entering at B and emerging at A is light passing through a medium with parallel surfaces (§ 279), and suffers no deviation. If we draw AB, cutting the axis at O, the triangles CAO, $C'BO$ are similar, and $\frac{CA}{C'B} = \frac{CO}{C'O}$. But $\frac{CA}{C'B}$, being the ratio of the radii, is constant for all parts of the surfaces, hence $\frac{CO}{C'O}$ must be constant, or all lines such as AB must cut the axis at one point O. O is the optical centre, and light passing through it is not deviated by the lens.

286. Geometrical Construction of Images.—For the geometrical construction of images formed by curved surfaces, it is convenient to use, in place of the waves themselves, lines perpendicular to the wave front, which represent the paths which the light follows, and are called *rays* of light. These rays, when perpendicular to a plane wave surface, are parallel, and an assemblage of such rays, limited by an aperture in a screen, is called a *beam*. When the rays are perpendicular to a spherical wave surface, they pass through the wave centre, and constitute a *pencil*.

A plane wave surface perpendicular to the axis of a lens is converted by the lens into a spherical wave surface with its centre at the principal focus. The rays perpendicular to the plane wave surface are parallel to the axis, and after emergence must all pass through the principal focus. Conversely, rays emanating from the principal focus emerge from the lens as rays parallel to the axis. Also, rays emanating from any focus must, after emerging from the lens, meet at the conjugate focus. Let L, Fig. 117, be a converging lens, and AB an object. Let O be the optical centre, and F the principal focus. Since all the rays from A must meet, after emerging from the lens, at the conjugate focus, which is the image of A, to find the position of the image it is only necessary to draw two such rays

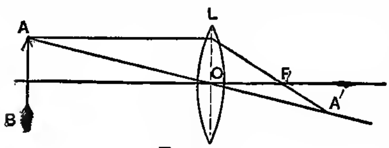

FIG. 117.

and find their intersection. The ray through the optical centre is not deviated, and the straight line AA' represents both the incident and emergent rays. The ray AL may be considered as one of a group parallel to the axis. All such rays must, after passing through the lens, pass through the principal focus. LA', passing through F, is therefore the emerging ray, and its intersection with AA' locates the image of A. Hence, to construct the image of a point, draw from the point two incident rays, and determine the corresponding emergent rays. The intersection of these will determine the image. The rays most convenient to use are the ray through the optical centre and the ray parallel to the axis or through the principal focus. Fig. 118 gives another example of an image determined by construction.

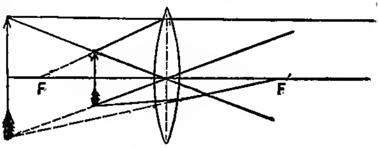

FIG. 118.

287. Thick Lenses.—When a lens is of considerable thickness, the formula derived in § 282 does not give the true position of the conjugate foci. A formula involving the thickness of the lens may be derived without difficulty, but for practical purposes it is usual to refer all measurements to two planes, called the *principal planes* of the lens. The determination of the position of these planes involves a discussion which does not come within the scope of this book.

288. Mirrors and Lenses of Large Aperture.—The equations derived in §§ 281, 282, are only approximations, applying with sufficient exactness to mirrors and lenses of small aperture. But for large apertures, terms containing the higher powers of x cannot be neglected, x will not disappear from the expression of p', and p' will, therefore, not have a definite value. In other words, the reflected or refracted wave is not spherical, and there is no one point l where the light will be concentrated. Surfaces may, however, be constructed which will, in certain particular cases, produce by reflection or refraction perfectly spherical waves. If we desire to find a surface such that light from L (Fig. 119) is concentrated by reflection at l, we remember that the sum $LB + Bl$ must be constant, and that this is a property of an ellipse with foci at L and l. If the ellipse be constructed and revolved about Ll as an axis, it will generate a surface which will have the required property. If one of the points L be removed to an infinite distance, the corresponding wave becomes a plane perpendicular to Ll, and we must have $LB + BC$ (Fig. 120) constant, a property of the parabola. A parabolic mirror will therefore concentrate at its focus incident light moving in paths parallel to its axis, or will reflect incident light diverging from its focus in plane waves perpendicular to its axis.

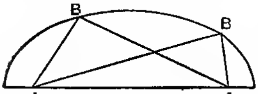

FIG. 119.

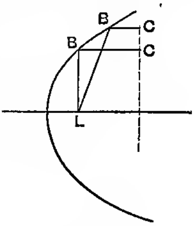

FIG. 120.

Mirrors and lenses having surfaces which are not spherical are seldom made because of mechanical difficulties of construction. It becomes necessary, therefore, to consider how the disadvantages arising from the use of spherical surfaces of large aperture for reflecting or refracting light may be avoided or reduced.

We will consider first the case of a spherical mirror. It was shown above that light from one focus of an ellipsoid is reflected from the ellipsoidal surface in perfectly spherical waves concentric with the other focus. Let Fig. 121 represent a plane section through the axis of an ellipsoid, and Fca a small incident pencil of light proceeding from the focus F. $F'ac$ is a section of the reflected pencil. It is a property of the ellipse that the normals to the curve bisect the angles formed by lines to the two foci. The normal ae bisects the angle FaF', and hence in the triangle FaF' we have $\dfrac{Fa}{F'a} = \dfrac{Fe}{F'e}$.

If a' move toward c, $F'a$ increases and Fa diminishes. Hence, from the above proportion, $F'e$ must increase and Fe diminish; or, the successive normals as we approach the minor axis cut the major axis in points successively nearer the centre of the ellipse. The normals produced will therefore meet each other at n beyond the axis. If ac be taken small enough it may be considered the arc of a circle of which an, cn are radii and n the centre. It is therefore a meridian section of an element of a spherical surface of which Fn is an axis.

Sections of wave surfaces reflected from the ellipsoid have their centre at F', and are also sections of wave surfaces reflected from the elementary spherical surface. Evidently the same would be true for any other meridian section passing

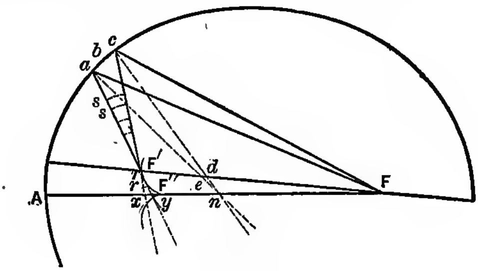

FIG. 121.

through FA of the sphere of which the elementary surface forms a part, and the form of the wave surfaces may be conceived by supposing the whole figure to revolve about FA as an axis. The arc ac describes a zone of the sphere, s, s, r, r, describe wave surfaces, and F' describes a circumference having its centre on FA. The wave surfaces are portions of the surfaces of curved tubes of which the axis is the arc described by the point F'. The line described by F' is a *focal line*, and all the light from the zone described by ac passes through it, or does so very approximately. If ac be taken nearer to A on the sphere, F' approaches the axis along the curve $F'F''$ and finally coincides with F'', the focus conjugate to F. $F'F''$ is a ***caustic curve***, which, when the figure revolves about the axis AF, describes a ***caustic surface***. It will be noted that all the light from the zone described by ac passes through the axis AF between the points x and y. The light coming from F and reflected from a small portion of the spherical surface around b, the middle point of ac, is then concentrated first in a line through F' at right angles to the paper, and again into the line xy in the plane of the paper. Nowhere is it concentrated into a point. A line drawn through b and the middle of the focal line through F' is the axis of the reflected pencil. It will intersect the axis of the mirror between x and y. If a plane be passed through the point of intersection perpendicular to the axis of the pencil, its intersection with the pencil will be like an elongated figure 8, which may be considered as a focal line at right angles to the axis of the pencil, and in the plane of the paper, and therefore at right angles to the focal line through F'. Between these two focal lines there is a section of least area, nearly circular, which is the nearest approach to an image of F produced by an oblique incidence such as we have been considering.

If refraction instead of reflection had taken place at ac, a result very similar would have been obtained for the refracted pencil. This failure of spherical reflecting or refracting surfaces to bring the light exactly to a focus is called ***spherical aberration***. In order to obtain a sharp focus, therefore, if only a single spherical surface be employed, the light must be confined within narrow limits of normal incidence. When reflection or refraction takes place at two or more surfaces in succession, the aberration of one may be made to partially correct the aberration of the other. For instance, when the waves incident upon a double convex lens are plane, the emerging waves are most nearly spherical when the radius of the second surface is six times that of the first. Two or more lenses may be so constructed and combined as to give, for sources of light at a certain distance, almost perfectly spherical emerging waves. Such combinations are called *aplanatic*. The same term is applied to single surfaces so formed as to give by reflection or refraction truly spherical waves.

SIMPLE OPTICAL INSTRUMENTS.

289. **The Camera Obscura.** — If a converging lens be placed in an opening in the window-shutter of a darkened room, well-defined images of external objects will be formed upon a screen placed at a suitable distance. This constitutes a *camera obscura*. The photographer's camera is a box in one side of which is a lens so adjusted as to form an image of external objects on a plate on the opposite side. The relation deduced in § 284 serves to determine the size of the image which a given lens will produce, or the focal length of a lens necessary to produce an image of a certain size.

290. **The Eye as an Optical Instrument.** — The eye, as may be seen from Fig. 122, which represents a section by a horizontal plane, is a camera obscura. *a* is a transparent membrane called the cornea, behind which is a watery fluid called the aqueous humor, filling the space between the cornea and the crystalline lens. Behind this is the vitreous humor, filling the entire posterior cavity of the eye. The aqueous humor, crystalline lens, and vitreous humor constitute a system of lenses, equivalent to a single lens of about two and a half centimetres focus, which produces a real inverted image of external objects upon a screen of nervous tissue called the retina, which lines the inner surface of the posterior half of the eyeball. The retina is an expansion of the optic nerve. The light that forms the image upon it excites the

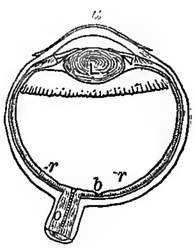

FIG. 122.

ends of the nerve, and, through the nerve-fibres leading to the brain, produces a mental impression, which, partly by the aid of the other senses, we have learned to interpret as the characteristics of the object the image of which produces the impression. For distinct vision the image must be sharply formed on the retina; but as an object approaches, its image recedes from a lens, and if, in the eye, there were no compensation, we could see distinctly objects only at one distance. The eye, however, adjusts itself to the varying distances of the object by changing the curvature of the front surface of the crystalline lens. There is a limit to this adjustment. For most eyes, an object nearer than fifteen centimetres does not have a distinct image on the retina.

We may here consider the means by which we estimate the distance and size of an object. The retina is not all equally sensitive. The depression at *b*, called the *yellow spot*, is much more sensitive than the other portions, and a minute area in the centre of that depression is much more sensitive than the rest of the yellow spot. That part of an image which falls on this small area is much more distinct than the other parts. How small this most sensitive area is, can be judged by carefully analyzing the effort to see distinctly the minute details of an object. For instance, in looking at the dot of an *i*, a change can be detected in the effort of the muscles that control the eyeball, when the attention is directed from the upper to the lower edge of the dot. The eye can then be directed with great precision to a very small object. The line joining the centre of the crystalline lens with the centre of the sensitive spot may be called the optic axis; and when the attention is directed to any particular point of an object, the eyeballs are turned by a muscular effort, until both the optic axes produced outward meet at the point. For objects at a moderate distance we have learned to associate a particular muscular effort with a particular distance, and our judgment of such distances depends mainly on this association. The angle between the optic axes when they meet at a point is called the *optic angle.* Our estimate of the size of an object is based on our judgment of its distance, together with the angle which the object subtends at the eye, called the *visual angle.* In Fig. 123, when *ab* is an object, and *l* the crystalline lens, α is the visual angle. It is plain that the size of the image on the retina is proportional to the visual angle. It is plain, too, that

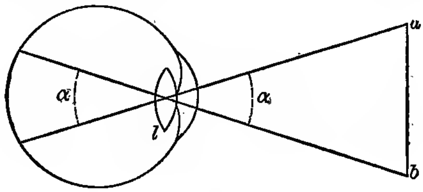

FIG. 123.

an object of twice the size, at twice the distance, would subtend the same visual angle and have an image of the same size as *ab.* Nevertheless, if we estimate its distance correctly we shall estimate its size as twice that of *ab*; but if in any way we are deceived as to its distance, and judge it to be less than it really is, we underestimate its size. Most persons underestimate heights, and hence underestimate the sizes of objects high above them. The visual angle is the *apparent size* of the object.

291. **Magnifying Power.**—To increase the apparent size of an object, and so improve our perception of its details, we must increase the visual angle. This can be done by bringing the object nearer the eye, but it is not always convenient or possible to bring an object near, and even with objects at hand there is a limit to the near approach, due to our inability to see distinctly very near objects. Certain optical instruments serve to increase the visual angle, and so improve our vision. Instruments for examining small objects, and increasing the

visual angle beyond that which the object subtends at the nearest point of distinct vision by the unaided eye, are called *microscopes*. Those used for observing a distant object and enlarging the visual angle under which it is seen at that distance are *telescopes*. In both cases the ratio of the visual angles, as the object is seen with the instrument, and without it, is the *magnifying power*.

292. The Magnifying Glass.—Fig. 124 shows how a converging lens may be employed to magnify small objects. The point *a* of an object just inside the principal focus *F* of the lens *A* is the origin of light-waves which, after passing through the lens, are changed to waves having a centre a' (§ 282) which, when the lens is properly adjusted, is at the distance of distinct vision. Waves coming from *b* enter the eye as though from b'. The object is therefore distinctly seen, but under a visual angle $a'Ob'$, while, to be seen distinctly by the unaided eye, it must be at the distance Oa'', when the angle subtended is $a''Ob''$. The ratio of these angles is very nearly that of Oa'' to OF. Hence the magnifying power is the ratio of the distance of distinct vision to the focal length of the lens.

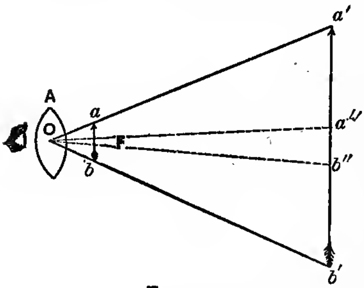

FIG. 124.

293. The Compound Microscope.—A still greater magnifying power may be obtained by first forming a real enlarged image of the object (§ 284) and using the magnifying glass upon the image, as shown in Fig. 125.

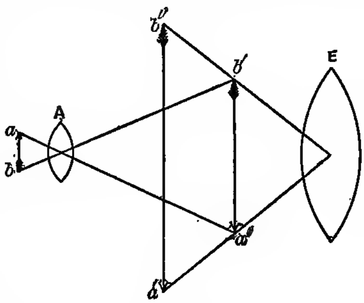

FIG. 125.

The lens A is called the *objective*, and E is called the *eye-lens* or *ocular*. As will be seen in § 310, both A and E often consist of combinations of lenses for the purpose of correcting aberration.

294. **Telescopes.**—If a lens or mirror be arranged to produce a real image of a distant object, either on a screen or in the air, we may observe the image at the distance of distinct vision when the visual angle for the object is enlarged in the ratio of the focal length of the lens to the distance of distinct vision. This will be plain from Fig. 126. Suppose the nearest

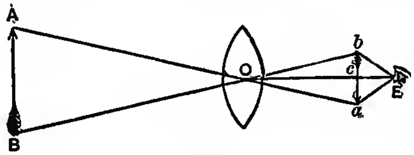

FIG. 126.

point from which the object can be observed by the naked eye to be the centre of the lens O. The visual angle is then $AOB = aOb$, while the visual angle for the image is aEb. Since these angles are always very small, we have

$$\frac{aEb}{aOb} = \frac{Oc}{Ec}$$

very nearly. But when AB is at a great distance, Oc is the focal length of the lens. By using a magnifying glass to observe the image, the magnifying power may be still further increased in the ratio of the distance of distinct vision to the focal length of the magnifying glass. The magnifying power of the combination is therefore the ratio of the focal length of the object-glass to the focal length of the eye-glass. A concave mirror may be substituted for the object-glass for producing the real image.

CHAPTER III.

VELOCITY OF LIGHT.

295. Velocity Determined from Eclipses of Jupiter's Moons.—Roemer, a Danish astronomer, was led to assume a progressive motion for light in order to explain some apparent irregularities in the motions of Jupiter's satellites. A few observations of one of Jupiter's moons are sufficient to determine the time of its eclipses for months in advance. If these observations be made when the earth and Jupiter are on the same side of the sun, and the time of an eclipse occurring about six months later, predicted from them, be compared with the observed time of that eclipse, it is found that the observed time is about $16\frac{2}{3}$ minutes later than the predicted time. This discrepancy is explained if it is assumed that light has a progressive motion and requires $16\frac{2}{3}$ minutes to cross the earth's orbit, for the distance of the earth from Jupiter in the second case is about the diameter of its orbit greater than in the first.

296. Aberration of the Fixed Stars.—The apparent direction of the light coming from a star to the earth, that is, the apparent direction of the star from the earth, is the resultant of the motion of the light and the motion of the earth. As the motion of the earth changes direction the apparent direction of the star will change also, and the amount of that change will depend on the relation between the velocity of light and the change in the velocity of the earth in its orbit, understanding by change of velocity change in direction as well as in amount. This apparent change in the position of the stars is called *aberra-*

tion. Knowing its amount corresponding to a known change in the earth's motion, we may compute the velocity of light. This method was first employed by Bradley.

297. Fizeau's Method.—Several methods have been employed for measuring the velocity of light by determining the time required for it to pass over a small distance on the earth's surface. In the form of experiment devised by Fizeau, a beam of light is allowed to pass out through a small hole in the shutter of a darkened room to a distant station where it is reflected back on itself. It returns through the opening and produces an image of the source. A toothed wheel is placed in front of the opening in such a position that, to pass out or back, the light must pass through the spaces between the teeth. If the wheel revolve slowly, as each space passes the opening in the shutter light will pass out, and returning from the distant station will enter through the space by which it made its exit. An image of the source will therefore be visible whenever a space passes the opening, and in consequence of the persistence of vision this image will appear continuous. Since it takes time for the light to go to the distant station and back, it is possible to give to the wheel such a velocity that when the light which passed out through a given space returns, it will find the adjacent tooth covering the opening, so that no image of the source can be seen. If the velocity of rotation be sufficiently increased, the image again comes into view when the light can enter through the space following that by which it emerged. A still further increase of velocity may cause a second extinction of the image. The experiment consists in determining accurately the velocities for which the several extinctions and reappearances of the image occur. A high degree of accuracy cannot be attained because the extinction of the image is not sudden. It disappears by a gradual fading away, and reappears by a gradual brightening. For quite a range of velocity the image cannot be seen at all.

298. Foucault's Method.—Foucault's method depends upon the use of the revolving mirror as a means of measuring a very small interval of time. Foucault's experiments were repeated with some modification by Michelson in 1879 and again in 1882. The general theory of the experiment may be understood from the following brief description. Let S (Fig. 127) be a narrow slit, m a mirror which may revolve about an axis in its own plane, L a lens, and m' a second mirror. Light from a source behind S passes through the slit, falls on m, is reflected, when m is in a suitable position, through the lens L,

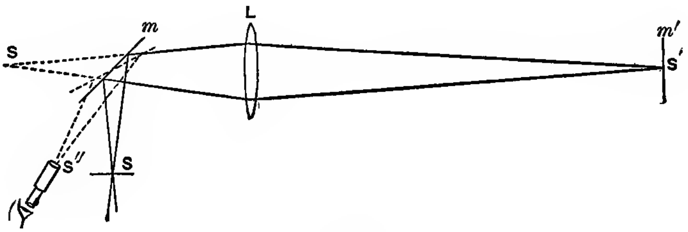

FIG. 127.

and forms an image at S'. S and S' are conjugate foci of the lens, and by so placing the lens that S shall be a little beyond the principal focus, S' may be removed to as great a distance as desired. The mirror m' is perpendicular to the axis of the lens, and at such a distance that the image S' falls upon its surface. It is evident that any light reflected back from m' through L will return to the conjugate focus S, whatever the position of the mirror m', so long as it sends the light in such a direction as to pass through L both going and returning. If now the mirror m be given a rapid rotation clockwise, light passing through L will return to find m in a changed position, and the image will be displaced from S to some point S'' to the left of S. Knowing the displacement SS'' and the number of rotations of the mirror per second, the time required for light to pass from m to S' and back is determined. The value of the velocity of light, as determined by Michelson in 1879, is 299,910, and in 1882, 299,853, kilometres per second.

299. Influence, upon the Velocity of Light, of the Motion of the Medium through which it Passes.—Fizeau showed by experiment in 1859 that a moving transparent body increases or diminishes the velocity of light passing through it, not by its own velocity, but by a fraction of its own velocity, expressed by $\frac{n^2 - 1}{n^2}$, where n is the index of refraction. This result was confirmed by experiments of Michelson and Morley in 1886. The result follows if we suppose the change of velocity of light in a medium to be due solely to change of density of the ether. Remembering that the velocity of propagation of wave motion in any medium is $v = \sqrt{\frac{E}{D}}$, and that the velocity in a medium of which the index of refraction is n, is $\frac{1}{n}$th as great as that in a vacuum, it may be seen at once that the density of the ether

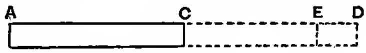

FIG. 128.

in such a medium must be n^2 times as great as that in a vacuum. In Fig. 128 let AC be a body, of which the index of refraction is n. Let the body move forward so as to occupy the position CD. The ether occupying the space CD and having a density 1 must in the body have a density n^2, and hence must occupy a space ED, which is $\frac{1}{n^2}$ times CD. The ether in AC must, therefore, move forward through a distance CE, while the body moves through a distance CD. But CE equals $CD \times \left(1 - \frac{1}{n^2}\right)$ or $CD \times \frac{n^2 - 1}{n^2}$. Hence the ratio of the velocity of the ether to the velocity of the body is $\frac{n^2 - 1}{n^2}$.

CHAPTER IV.

INTERFERENCE AND DIFFRACTION.

300. Interference of Light from Two Similar Sources.— It has already been shown that the disturbance propagated to any point from a luminous wave is the algebraic sum of the disturbances propagated from the various elements of the wave. The phenomena due to this composition of light-waves are called *interference* phenomena.

Let us consider the case in which two elements only are

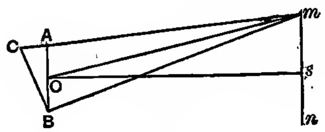

FIG. 129.

efficient in producing the disturbance. Let A and B (Fig. 129) represent two elements of the same wave surface separated by the very small distance AB. The disturbance at m, a point on a distant screen mn, parallel with AB, due to these two elements, is the resultant of the disturbances due to each separately. The light is supposed to be homogeneous, and its wave length is represented by λ.

When the distance $mB - mA$ equals $\frac{1}{2}\lambda$, or any odd multiple of $\frac{1}{2}\lambda$, there will be no disturbance at m. Take $mC = mB$, and draw BC. mCB is an isosceles triangle; but since AB is very small compared to Om, the angle at C may be taken as a right angle; the triangle ACB, therefore, is similar to Osm, and we have

$$\frac{AB}{AC} = \frac{Om}{sm} = \frac{Os}{sm} \text{ very nearly.}$$

Represent sm by x, Os by c, AB by b, AC by $n \times \frac{1}{2}\lambda$, where n is any number. Then we have

$$x = \frac{\frac{1}{2}\lambda cn}{b}. \qquad (115)$$

If n be any even whole number, the values of x given by this equation represent points on the screen mn at which the waves from A and B meet in the same phase and unite to produce light. If n be any odd whole number, the corresponding values of x represent points where the waves meet in opposite phases, and therefore produce darkness. It appears, therefore, that starting from s, for which $n = 0$, we shall have darkness at distances

$$\frac{\frac{1}{2}\lambda c}{b}, \frac{\frac{3}{2}\lambda c}{b}, \frac{\frac{5}{2}\lambda c}{b}, \text{ etc.,}$$

and light at distances

$$0, \frac{\lambda c}{b}, \frac{2\lambda c}{b}, \frac{3\lambda c}{b}, \text{ etc.}$$

From Eq. (115), we have

$$n = \frac{2bx}{c\lambda}.$$

Since $\frac{1}{2}n\lambda$ is the number of wave lengths that the wave front from B falls behind that from A, $\frac{1}{2}nT$, where T represents the period of one vibration, is the time that must elapse after the wave from A produces a certain displacement before that from B produces a similar displacement. The expression

$$\frac{2\pi\frac{1}{2}nT}{T} = n\pi$$

is, therefore, the difference in epoch of the two wave systems. Substituting $n\pi$ for ϵ in Eq. (9), we have

$$S = s + s_{\prime} = a(2 + 2\cos n\pi)^{\frac{1}{2}} \cos\left(\frac{2\pi t}{T} - \tan^{-1}\frac{\sin n\pi}{1 + \cos n\pi}\right).$$

Now the intensity of light for a vibration of any given period is proportional to the energy of the vibratory motion. It is therefore proportional to the square of the maximum velocity, and this is proportional to the square of the amplitude. To find the relative intensities of light at different points, we may suppose t in the second parenthesis above to have such a value as shall render the cosine unity, when

$$S = a(2 + 2\cos n\pi)^{\frac{1}{2}} = A$$

is the amplitude of the vibratory motion for any given value of n. Substituting for n its value and squaring, we have

$$A^2 = a^2\left(2 + 2\cos\frac{2bx}{c\lambda}\pi\right),$$

in which A^2 is proportional to the intensity of the illumination at distances x from s. When

$$\frac{2bx}{c\lambda}\pi = 0,$$

its cosine is 1, and A^2 is a maximum and equal to $4a^2$. As x increases A^2 diminishes, until

$$\frac{2bx}{c\lambda}\pi = \pi, \text{ in which case } A^2 = 0.$$

A^2 then increases until it becomes again a maximum, when

$$\frac{2bx}{c\lambda}\pi = 2\pi.$$

In short, if AB (Fig. 130) represent the line mn of Fig. 129, the ordinates to a sinuous curve like abc will represent the intensities of the light along that line.

The phenomena described above may be realized experimentally in several ways. Young admitted sunlight into a darkened room through a small hole in a window-shutter. It fell upon a screen in which were two small holes close together, and, on passing through these, was received upon a second screen. Light and dark bands were observed upon this screen, the distances of which from the central band were in accordance with theory.

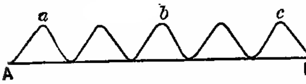

FIG. 130.

Fresnel received the light from a small luminous source upon two mirrors making a very large angle, as in Fig. 131. The light reflected from each mirror proceeded as though from the image of the source produced by that mirror. The reflected light, therefore, consisted of two wave systems, from two precisely similar sources A and B. Light and dark bands were formed in accordance with theory. In order that the experiment may be successfully repeated reflection must take place

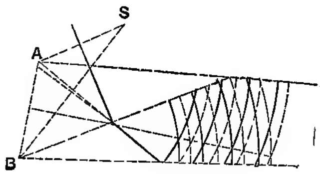

FIG. 131.

from the front surface of each mirror only, the angle made by the mirrors must be nearly 180°, and the reflecting surfaces must meet exactly at the vertex of the angle. Two similar sources of light may be obtained also by sending the light through a double prism, as shown in Fig. 132. Light from A proceeds after passing through the prism as from the two virtual images a and a'.

FIG. 132.

A divided lens, Fig. 133, serves the same purpose. The light from A is concentrated in two real images a and a', from which proceed two wave systems as in the previous cases. What are really seen in these cases, when the source of light is white, are iris-colored bands instead of bands of light and darkness merely.

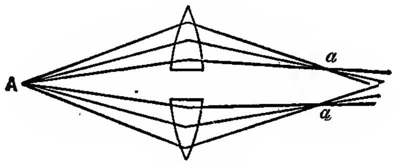

FIG. 133.

When the light is monochromatic, the bands are simply alternations of light and darkness, the distances between them being greatest for red light, and least for blue. From Eq. (115) it appears that, other things being equal, x varies with λ, hence we must conclude that the greater distance between the bands indicates a greater wave length; that is, that the wave length of red light is greater than that of blue.

301. Measurement of Wave Lengths.—Data may be obtained from any of the above experiments for the determination of the wave length of light. From Eq. (115) we have

$$\lambda = \frac{2bx}{cn},$$

where c, b, and x are distances to be measured. The distance x is the distance from s to a point m, the centre of a light band, and n equals twice the number of dark bands between s and m. It is not necessary to consider the details of the apparatus, and the adjustments necessary for making these measurements. It is sufficient to show, in a general way, how the distance x can be measured. Instead of a screen, a lens or combination of lenses, called a positive eyepiece, is placed in the path of the light, and the observer looks through it towards the luminous source. This eyepiece has a spider-line stretched in front of it, which is seen magnified when the bands are observed, and lens and spider-line are arranged to be moved laterally by a micrometer screw. By this movement the spider-line may be brought to coincide with the bands in succession, and the distances measured by the number of revolutions of the screw. Better methods than this of measuring wave lengths will be found described in § 306.

302. Interference from Thin Films.—Thin films of transparent substances, such as the wall of a soap-bubble or a film of oil on water, present interference phenomena when seen in a strong light, due to the interference of waves reflected from the two surfaces of the film. Let AA, BB (Fig. 134) be the surfaces of a transparent film. Light falling on AA is partly reflected and partly transmitted. The reflection at the upper surface takes place with change of sign (§ 246). The light entering the film is partly reflected at the lower surface without change of sign, and returning partly emerges at the upper surface. It is there compounded with the wave at that moment reflected. Let us suppose

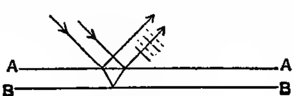

FIG. 134.

the light homogeneous, and the thickness of the film such that the time occupied by the light in going through it and returning is the time of one complete vibration. The returning wave will be in the same phase as the one at that moment entering, and, therefore, opposite in phase to the wave then reflected. The reflected and emerging waves destroy each other, or would do so if their amplitudes were equal, and the result is that, apparently, no light is reflected. If the light falling on the film be white light, any one of its constituents will be suppressed when the time occupied in going through the film and returning is the period of one vibration, or any whole number of such periods, of that constituent. The remaining constituents produce a tint which is the apparent color of the film.

Similar phenomena are produced by the interference of that portion of the incident light which is transmitted directly through the film, with that portion which is transmitted after undergoing an even number of internal reflections. Since these reflections occur without change of sign, the thickness of the film for which the reflected light is a minimum is that for which the transmitted light is a maximum.

Newton was the first to study these phenomena. He placed a plane glass plate upon a convex lens of long radius, and thus formed between the two a film of air, the thickness of which at any point could be determined when the radius of the sphere and the distance from the point of contact were known. With this arrangement Newton found a black spot at the point of contact, and surrounding this, when white light was used, rings of different colors. When homogeneous light was used, the rings were alternately light and dark. Let ab (Fig. 135), be the radius of the first dark ring, and denote it by d. The thickness $bc = ef$, which may be denoted by x, is

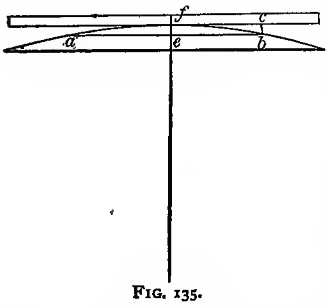

FIG. 135.

$$x = \frac{d^2}{2r - x}.$$

Since x is very small in comparison with $2r$, this becomes

$$x = \frac{d^2}{2r}.$$

This distance for the first dark ring, when the incident light is normal to the plate, is equal to half the wave length of the light experimented upon. Newton found the thickness for the first dark ring $\frac{2}{178000}$ inches, which corresponds to a wave length of about $\frac{1}{44500}$ inches, or 0.00057 *mm*. This method affords a means of measuring the wave lengths of light, or, if the wave lengths be known, we may determine the thickness of a film at any point.

303. Effects Produced by Narrow Apertures.—It has been seen (§ 274), that cutting off a portion of a light-wave by means of screens, thus leaving a narrow aperture for the passage of the light, prevents the interference which confines the light to straight lines, and gives rise to a luminous disturbance within the geometrical shadow. This phenomenon is called *diffraction*. Let us consider the aperture perpendicular to the plane of the paper, and an approaching plane wave parallel to the plane of the aperture. Let AB (Fig. 136) represent the aperture, and mn one position of the approaching wave. To determine the effect at any point we must consider the elementary waves proceeding from the various points of the wave front lying between A and B. First consider the point P on the perpendicular to AB at its middle point. AB is so small that the distances from P to each point of AB may be regarded as equal, or the time of passage of the light from each point of AB to P may be made

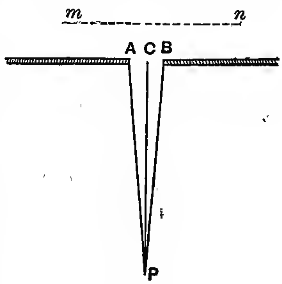

FIG. 136.

the same, by placing a converging lens of proper focus between AB and P. Then all the elementary waves from points of AB meet at P in the same phase, and the point P is illuminated. Now consider a second point, P', in an oblique direction from C, Fig. 137, and suppose the obliquity such that the time of passage from B to P' is half a vibration period less than the time of passage from C to P', and a whole vibration period less than the time of passage from A to P'. Plainly the elementary waves from B and C will meet at P' in opposite phases, and every wave from a point between B and C will meet at P a wave in the opposite phase from some point between C and A. The point P' is, therefore, not illuminated.

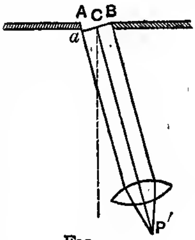

FIG. 137.

Suppose another point, P'' (Fig. 138), still further from P, such that AB may be divided into three equal parts, each of which is half a wave length nearer P'' than the adjacent part. It is plain that the two parts Bc and ca will annul each other's effects at P'', but that the odd part Aa will furnish light. At a greater obliquity, AB may be divided into four parts, the distances of which from the point, taken in succession, differ by half a wave length. There being an even number of these parts, the sum of their effects at the point will be zero. Now let us suppose the point P to occupy successively all positions to the right or left of the normal. While the line joining P with the middle of the aperture is only slightly oblique, the elementary waves meet at P in nearly the same phase, and the loss of light is small. As P approaches P' (Fig. 137), more and more of the waves meet in opposite phases, the light grows rapidly less, and at P' becomes zero. Going beyond P'

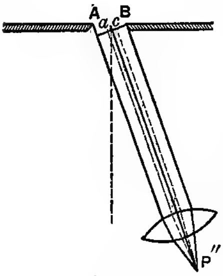

FIG. 138.

the two parts that annul each other's effects no longer occupy the whole space AB, some of the points of the aperture send to P waves that are not neutralized, and the light reappears, giving a second maximum, much less than the first in intensity. Beyond this the light diminishes rapidly in intensity until a point is reached where the paths differing by half a wave length divide AB into four parts, when the light is again zero. Theoretically, maximum and minimum values alternate in this way, to an indefinite distance, but the successive maxima decrease so rapidly that, in reality, only a few bands can be seen.

304. Effect of a Narrow Screen in the Path of the Light.—It can be shown that the effect of a narrow screen is the complement of that of a narrow aperture; that is, where a narrow aperture gives light, a screen produces darkness. Let mn (Fig. 139) be a plane wave and AB a surface on which the light falls. If no obstacle intervene, the surface AB will be equally illuminated. The illumination at any point C is the sum of the effects of all parts of the wave mn. Let the effects due to the part of the wave op be represented

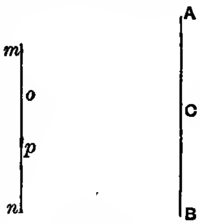

FIG. 139.

by I and that due to all the rest of the wave by I'. Then the illumination at C is $I + I'$, equal to the general illumination on the surface. Let us now suppose mn to be a screen and po a narrow aperature in it. If the illumination at C remain unchanged, it must be that the parts mo and pn of the wave had no effect, and if, for the screen with the narrow aperture, we substitute a narrow screen at op, there will be darkness at C. If, however, a dark band fall at C when op is an aperture, a screen at op will not cut off the light from C. That is, if C be illuminated when op is an aperture, it will be in darkness when op is a screen, and if it be in darkness when op is an aperture, it will be illuminated when op is a screen.

305. Diffraction Gratings.—Let AB (Fig. 140) be a screen having several narrow rectangular apertures parallel and equidistant. Such a screen is called *grating*. Let the approaching waves, moving in the direction of the arrow, be plane and parallel to AB. Draw the parallel lines ab, cd, etc., at such an angle that the distance from the centre of a to the foot of the perpendicular let fall from the centre of the adjacent opening on ab shall be equal to some definite wave length of light. It is evident that an will contain an exact whole number of wave lengths, co one wave length less, etc. The line mn is, therefore, tangent to the fronts of a series of elementary waves which are in the same phase, and may be considered as a plane wave, which, if it were received on a converging lens, would be concentrated to a focus. If the obliquity of the lines be increased until ae equals 2λ, 3λ, etc., the result will be the same. Let us, however, suppose that ae is not an exact multiple of a wave-length, but some fractional part of a wave length, $\frac{99}{100}\lambda$ for example. Let m be the fifty-first opening counting from a; then an will be $\frac{99}{100}\lambda \times 50 = 49.5\lambda$. Hence the wave from the first opening will be in the opposite phase to that from the fifty-first. So the wave from the second opening will be in the opposite phase to that from the fifty-second, etc. If there were one hundred openings in the screen, the second fifty would exactly neutralize the effect of the first fifty in the direction assumed. Light is found, therefore, only in directions given by

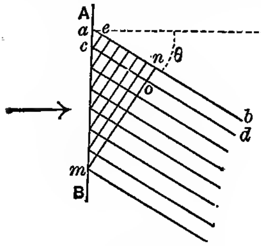

FIG. 140.

$$\sin\theta = \frac{n\lambda}{d}, \qquad (116)$$

where n is a whole number, θ the angle between the direction of the light and the normal to the grating, and d the distance from centre to centre of the openings, usually called an *element* of the grating. Gratings are made by ruling lines on glass at the rate of some thousands to the centimetre. The rulings may also be made on the polished surface of speculum metal, and the same effects as described above are produced by reflection from its surface. Since the number of lines on one of these gratings is several thousands, it is seen that the direction of the light is closely confined to the direction given by the formula, or, in other words, light of only one wave length is found in any one direction. If white light, or any light consisting of waves of various lengths, fall on the grating, the light corresponding to different wave lengths will make different angles with AC, that is, the light is separated into its several constituents and produces a *pure spectrum*. Since different values of n will give different values of θ for each value of λ, it is plain that there will be several spectra corresponding to the several values of n. When n equals 1 the spectrum is of the *first order;* when n equals 2 the spectrum is of the *second order*, etc. The grating furnishes the most accurate and at the same time the most simple method of determining the wave lengths of light. Knowing the width of an element of the grating it is only necessary to measure θ for any given kind of light.

In this discussion it has been assumed that the light was normal to the surface of the grating. This need not be the case. Let AB (Fig. 141) be the intersection with the paper of a reflecting grating supposed perpendicular to it, mn an approaching wave front also perpendicular to the paper, and $m''n''$ the reflected wave front constructed as in § 278. The line $m''n''$ is a tangent to all the elementary waves that originate in the surface AB in consequence of disturbances produced by the passage of the wave

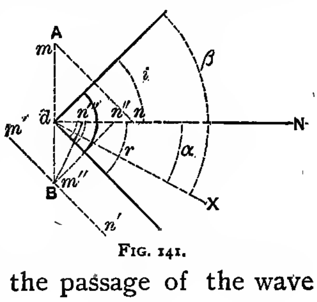

FIG. 141.

$m'n'$. The surface AB consists of a number of narrow, equidistant, reflecting surfaces separated by roughened channels. If the reflecting surfaces be considered infinitely narrow, each of them will be the centre of a system of waves due to the successive incident waves similar to mn which fall upon them. Since the number of the elements of the grating is finite there will be a finite number of such wave systems. In the diagram one of these systems is represented about the centre d. Let us represent by a, b, c, d, etc., the centres of these systems, such that the distances $m''a$, ab, bc, cd, etc., are elements of the grating.

Let us suppose the wave systems all represented, and draw $m''n'''$ tangent to the wave front of which the centre is a, and which is one wave length behind the wave to which $m''n''$ is tangent. The line $m''n'''$ will be also tangent to waves of the systems of which b, c, d, etc., are the centres, and which are respectively two, three, four, etc., wave lengths behind the wave to which $m''\,n''$ is tangent. These elementary waves, differing by successive periods, are all in the same phase, and $m''n'''$ may, therefore, be considered as constituting a plane wave front in which light of one particular wave length is propagated in the direction dx. Represent by i the angle of incidence, by r the angle of reflection, by α the angle between the normal to the grating and the path of the diffracted light. Then i equals r, and if $m''a$ equal s, the radius of the elementary wave having its centre at a, and tangent to $m''n''$, is $s \sin i$, and of the elementary wave having the same centre, and tangent to $m''n'''$, is $s \sin \alpha$. Hence, by hypothesis, we have $s \sin i - s \sin \alpha = \lambda$.

Let us designate by β the angle between the path of the incident and that of the diffracted light, and by θ the angle between the path of the reflected and that of the diffracted light. If the grating be turned so that the path of the reflected light coincides with dx, its normal will turn through the angle $\frac{\theta}{2}$ and will bisect the angle β. Hence we have $i = \frac{\beta}{2} + \frac{\theta}{2}$, and $\alpha = \frac{\beta}{2} - \frac{\theta}{2}$. Substituting these values in the equation for λ we obtain

$$\lambda = 2s \cos \frac{\beta}{2} \sin \frac{\theta}{2}. \tag{117}$$

Hitherto the spaces from which the elementary waves proceed have been considered infinitely narrow, so that only one system of waves from each space need be considered. In practice, these spaces must have some width, and it may happen that the waves from two parts of the same space may cancel each other. Let the openings, Fig. 142, be equal in width to the opaque spaces, and let the direction am be taken such that ae equals 2λ. Then ae' equals $\frac{1}{2}\lambda$, or the waves from one half of each opening are opposite in phase to those from the other half, and there can be no light in the direction am. In general, if d equal the width of the opening, there will be interference and light will be destroyed in that direction for which

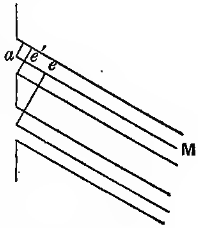

FIG. 142.

$\sin \theta = \frac{n\lambda}{d}$, if the incident light be normal to the grating. Let f represent the width of the opaque space. Then $d + f = s$, and light occurs in the direction given by $\sin \theta = \frac{n\lambda}{d+f}$, provided that the value of θ given by this equation does not satisfy the first equation also.

If d equal f, we have

$$\sin \theta = \frac{n\lambda}{d+f} = \frac{n\lambda}{2d}.$$

When n is even, $\sin \theta$ becomes

$$\frac{2\lambda}{2d} = \frac{\lambda}{d}; \quad \frac{4\lambda}{2d} = \frac{2\lambda}{d}, \quad \text{etc.},$$

and satisfies the equation

$$\sin \theta = \frac{n\lambda}{d},$$

which expresses the condition under which light is all destroyed. Hence in this case all the spectra of even orders fail. Moreover, the spectra after the first are not brilliant. When f equals $2d$ the spectrum of the third order fails.

It may be shown that whatever be the relative widths of the transparent and opaque spaces, one may be substituted for the other without altering the result. In Fig. 143 let ac represent an opening and cd an opaque portion. Let us assume that cd equals $3ac$, and let ab be the path of the diffracted light giving the spectrum of the first order; then we have $ae = \lambda$ and $ae' = \frac{1}{4}\lambda$. Now let ac become the opaque portion and cd the opening. We will then have $ih = \frac{1}{2}\lambda$. Each of the elementary waves from points between c and i will be half a wave length behind a corresponding wave from some point between d and j, so that the waves coming from ci and dj annul one another, and ij is the only efficient portion of the opening cd. This portion ij is equal to the former opening ac. Since the effect of the grating is that of one opening multiplied by the number of openings, it is plain that in this case it is indifferent whether the openings are of the width ac or cd.

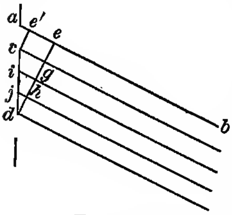

Fig. 143.

306. Measurement of Wave Lengths.—To realize practically the conditions assumed in the theoretical discussion of the last section, some accessory apparatus is required. It has been assumed that the wave incident upon the grating was plane. Such a wave would proceed from a luminous point or line at an infinite distance. In practice it may be obtained by illuminating a very narrow slit, taking it as the source of light, and placing it in the principal focal plane of a well-corrected converging lens. The plane wave thus obtained passes through the grating, or is reflected from it, and is received on a second lens similar to the first, which gives an image either on a screen or in front of an eyepiece, where it is viewed by the eye. The general construction of the apparatus may be inferred from Fig. 144. It is called the *spectrometer*.

A is a tube carrying at its outer end the slit and at its inner end the lens, called a *collimating* lens. CD is a horizontal graduated circle, at the centre of which is a table on which the grating is mounted, and so adjusted that the axis of the circle lies in its plane and parallel to its lines. In using a reflecting grating the collimating and observing telescopes may be fixed at a constant angle with each other which may be determined once for all in making the adjustments of the instrument. This angle is the angle β of § 305. To determine this angle the grating is turned until light thrown through the observing telescope upon the grating is reflected back on itself. The position of the graduated circle is then read. The difference between this reading and the reading when the grating is in such a position that the reflected image of the slit is seen in the telescope is the angle $\frac{\beta}{2}$. If the grating be now turned until the light of which the wave length is required is observed, the angle through which it is turned from its last position is the angle $\frac{\theta}{2}$. If the width of an element of the grating be known, these measurements substituted in Eq. 117 give the value of λ.

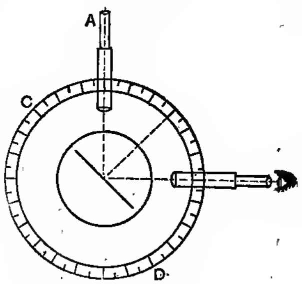

FIG. 144.

Wave lengths are generally given in terms of a unit called a *tenth metre;* that is, 1 metre $\times$ 10^{-10}. The wave lengths of the visible spectrum lie between 7500 and 3900 tenth metres. Langley has found in the lunar radiations wave lengths as long as 170,000 tenth metres, and Rowland has obtained photographs of the solar spectrum in which are lines representing wave lengths of about 3000 tenth metres.

Instead of the arrangement which has been described, Rowland has devised a grating ruled on a concave surface, and is thus enabled to dispense with the collimating lens and the telescope.

307. Phenomena due to Diffraction.—The colors exhibited by mother-of-pearl are due to diffraction effects produced by the striated surface. Luminous rings are sometimes seen closely surrounding the sun or moon, due to small globules of vapor or particles of ice in the upper atmosphere. Similar rings may be seen by looking at a small luminous source through a plate of glass strewn with lycopodium powder.

CHAPTER V.

DISPERSION.

308. Dispersion.—When white light falls upon a prism of any refracting medium, it is not only deviated from its course but separated into a number of colored lights, constituting an image called a spectrum. These merge imperceptibly from one into another, but there are six markedly different colors: red, orange, yellow, green, blue, and violet. Red is the least and violet the most deviated from the original course of the light. Newton showed by the recomposition of these colors by means of another prism, by a converging lens, and by causing a disk formed of colored sectors to revolve rapidly, that these colors are constituents of white light, and are separated by the prism because of their different refrangibilities. To arrive at a clear understanding of the formation of this spectrum, let us suppose first a small source of homogeneous light L (Fig. 145). If this light fall on a converging lens from a point

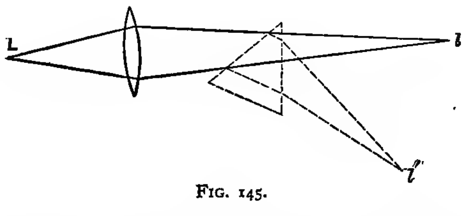

FIG. 145.

at a distance from it a little greater than that of the principal focus, a distinct image of the source will be formed at the distant conjugate focus l. If now a prism be placed in the path of the light, it will, if placed so as to give the minimum deviation, merely deviate the light without interfering with the sharpness of the image, which will now be formed at l' instead of at l. If the source L give two or three kinds of light, the lens may be so constructed as to produce a single sharp image at l of the same color as the source, but when the prism is introduced the lights of different colors will be differently deviated and two or three distinct images will be found near l'. If there be many such images, some may overlap, and if there be a great number of kinds of light varying progressively in refrangibility, there will be a great number of overlapping images constituting a continuous spectrum.

309. **Dispersive Power.**—It is found that prisms of different substances giving the same mean deviation of the light deviate the light of different colors differently, and so produce a longer or shorter spectrum. The ratio of the difference between the deviations of the extremities of the spectrum to the mean deviation may be called the *dispersive power* of the substance. Thus if d', d'' represent the extreme deviations, and d the mean deviation, the dispersive power is $\frac{d' - d''}{d}$.

In § 279 we find the equation $\frac{\sin Arx}{\sin Axr} = \mu$, and referring to Fig. 108 we may set $\sin Arx = \sin (Axr + xAr)$. From the discussion of § 279 it appears that when the prism is in the position of minimum deviation, the angle Axr equals half the refracting angle of the prism, or $\frac{A}{2}$, and the angle xAr equals half the deviation, or $\frac{d}{2}$. Hence we obtain

$$\mu = \frac{\sin \frac{A + d}{2}}{\sin \frac{A}{2}}; \tag{118}$$

or when A is small, $\mu = \dfrac{\frac{A+d}{2}}{\frac{A}{2}}$;

from which $d = A(\mu - 1)$.

Hence we obtain

$$\frac{d' - d''}{d} = \frac{A(\mu' - 1) - A(\mu'' - 1)}{A(\mu - 1)} = \frac{\mu' - \mu''}{\mu - 1},$$

where μ' and μ'' are the refractive indices for the extreme colors, and μ the index for the middle of the spectrum.

310. Achromatism.—If in Newton's experiment of recomposition of white light by the reversed prism the second prism be of higher dispersive power than the first, and of such an angle as to effect as far as possible the recomposition, the light will not be restored to its original direction, but will still be deviated, and we shall have deviation without dispersion. This is a most important fact in the construction of optical instruments. The dispersion of light by lenses, called *chromatic aberration*, was a serious evil in the early optical instruments, and Newton, who did not think it possible to prevent the dispersion, was led to the construction of reflecting telescopes to remedy the evil. It is plain, however, from what has been said above, that in a combination of two lenses of different kinds of glass, one converging and the other diverging, one may correct the dispersion of the other within certain limits, while the combination still acts as a converging lens forming real images of objects. Fig. 146 shows how this principle is applied to the

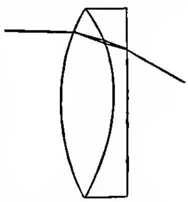

FIG. 146.

correction of chromatic aberration in the object-glasses of telescopes.

Thus far nothing has been said of the relative separation of the different colors of the spectrum by refraction by different substances. Suppose two prisms of different substances to have such refracting angles that the spectra produced are of the same length. If these two spectra be superposed, the extreme colors may be made to coincide, but the intermediate colors do not coincide at the same time for any two substances of which lenses can be made. Perfect achromatism by means of lenses of two substances is therefore impossible. In practice it is usual to construct an achromatic combination to superpose, not the extreme colors, but those that have most to do with the brilliancy of the image.

The indistinctness due to chromatic aberration, existing even in the compound objective, may be much diminished by a proper disposition of the lenses of the eyepiece. Fig. 147 shows the *negative* or Huyghens eyepiece.

Let A be the objective of a telescope or microscope. A

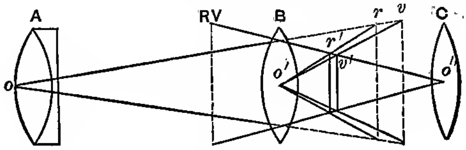

FIG. 147.

point situated on the secondary axis ov would, if the objective were a single lens, have images on that axis, the violet nearest and the red farthest from the lens. If the lens could be perfectly corrected, these images would all coincide. By making the lens a little over-corrected, the violet may be made to fall beyond the red. Suppose r and v to be the images. B and C are the two lenses of the Huyghens eyepiece. B is called the *field-lens*, and is three times the focal length of C. It is placed between the objective and its focal plane, and therefore prevents the formation of the images rv, but will form images at $r'v'$ on the secondary axes $o'r$, $o'v$. If everything is properly proportioned, $r'v'$ will fall on the secondary axis $o''R$ of the eye-lens C at such relative distances as to produce *one* virtual image at RV. It will be noted that the image r' is smaller than would have been formed by the objective. The magnifying power of the instrument is therefore less than it would be if the lens C were used alone as the eyepiece. This loss of magnifying power is more than counterbalanced by the increased distinctness.

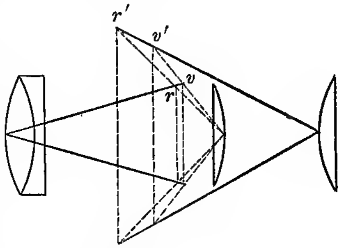

FIG. 148.

Fig. 148 shows the *Ramsden* or *positive* eyepiece. The aid it gives in correcting the residual errors of the objective is evident from the figure.

311. The Rainbow.—The rainbow is due to refraction and dispersion of sunlight by drops of rain. The complete theory of the rainbow is too abstruse to be given here, but a partial explanation may be given. Let O, Fig. 149, represent a drop of water, and SA the paths of the incident light from the sun. The light enters the drop, suffers refraction on entrance, is reflected from the interior surface near B, and emerges near C, as a wave of double curvature of which mn may be taken as the section. Of this wave the part near p, the point of inflection, gives the maximum effect at a distant point, and if the eye be placed in the prolongation of the line CE perpendicular to the wave surface, light will be perceived, but at a very little distance above or below CE there will be darkness. The direction CE is very nearly that of the minimum deviation

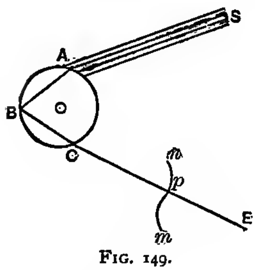

FIG. 149.

produced by the drop with one internal reflection. It is also the direction in which the angle of emergence equals the angle of incidence. The direction *CE* corresponds to the minimum deviation for only one kind of light. If this be red light, the yellow will be more deviated, and the blue still more. To see these colors the eye must be higher up, or the drop lower down. If the eye remain stationary, other drops below *O* will send to it the yellow and blue, and other colors of the spectrum. Since this effect depends only on the angle between the directions *SA* and *CE*, it is clear that a similar effect will be received by the eye at *E* from all drops lying on the cone swept out by the revolution of the line *CE* and all similar lines drawn to the drops above and below the drop *O*, about an axis drawn through the sun and the eye, and hence parallel to *SA*. This cone will trace out the primary rainbow having the red on the outer and the blue on the inner edge. The secondary bow, which is fainter, and appears outside the primary, is produced by two reflections and refractions as shown in Fig. 150.

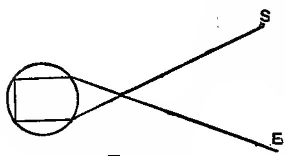

FIG. 150.

312. **The Solar Spectrum.**—As has been seen (§ 308) solar light when refracted by a prism gives in general a continuous spectrum. Wollaston, in 1802, was the first to observe that when solar light is received upon a prism through a very narrow opening at a considerable distance, dark lines are seen crossing the otherwise continuous spectrum. Later, in 1814–15, Fraunhofer studied these lines, and mapped about 600 of them. That these may be well observed in the prismatic spectrum it is important that the apparatus should be so constructed as to avoid as far as possible spherical and chromatic aberrations. The slit must be very narrow, so that its images may overlap as little as possible. The most important condition for avoiding spherical aberration is that the waves reaching the prism should be plane waves, since all others are distorted by refraction at a plane sur-

face. Fig. 151 shows the disposition of the essential parts of the apparatus known as the *spectroscope*. *S* is the slit, which may be considered as the source of light. *C* is an achromatic lens, called a *collimating* lens, so placed that *S* is in its principal focus. The waves emerging from it will then be plane. These will be deviated by the prism, and the waves representing the different colors will be separated, so that after passing through the second lens *O* these different colors will each give a separate

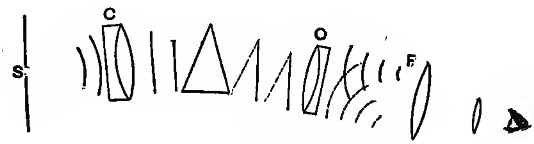

Fig. 151.

image. These images may be received upon a screen, or observed by means of an eyepiece. Sometimes a series of prisms is used to cause a wider separation of the different images.

If the images at *F* be received on a sensitive photographic plate, it will be found that the image extends far beyond the visible spectrum in the direction of greater refrangibility, and a thermopile or bolometer will show that it also extends a long distance in the opposite direction beyond the visible red. The solar radiations, therefore, do not all have the power of exciting vision. Much the larger part of the solar beam manifests its existence only by other effects. It will be shown that, physically, the various constituents into which white light is separated by the prism differ essentially only in wave length.

313. Spectrum Analysis.—If, in place of sunlight, the light of a lamp or of any incandescent solid, such as the lime of the oxyhydrogen light or the carbons of the electric lamp, illuminate the slit, a continuous spectrum like that produced by sunlight is seen, but the black lines are absent. Solids and liquids give in general only continuous spectra. Gases, however, when incandescent give continuous spectra only very rarely. Their spectra are bright lines which are distinct and separate images of the slit. The number and position of these lines differ with each gas employed. Hence, if a mixture of several gases not in chemical combination be heated to incandescence, the spectral lines belonging to each constituent, provided all be present in sufficient quantity, will be found in the resultant spectrum. Such a spectrum will therefore serve to identify the constituents of a mixture of unknown composition. Many chemical compounds are decomposed into their elements, and the elements are rendered gaseous at the temperature necessary for incandescence. In that case the spectrum given is the combined spectra of the elements. A compound gas that does not suffer dissociation at incandescence gives its own spectrum, which is, in general, totally different from the spectra of its elements.

The appearance of a gaseous spectrum depends in some degree on the density of the gas. When the gas is sufficiently compressed, the lines become broader and lose their sharply defined edges, and if the compression be still further increased the lines may widen until they overlap, and form a continuous spectrum. Some of the dark lines of the solar spectrum are found to coincide in position with the bright lines of certain elements. This coincidence is absolute with the most perfect instruments at our command, and not only so, but if the bright lines of the element differ in brilliancy the corresponding dark lines of the solar spectrum differ similarly in darkness.

The close coincidence of some of these lines was noted as early as 1822 by Sir John Herschel, but the absolute coincidence was demonstrated by Kirchhoff, who also pointed out its significance. Placing the flame of a spirit lamp with a salted wick in the path of the solar beam which illuminated the slit of his spectroscope, Kirchhoff found the two dark lines corresponding in position to the two bright lines of sodium to become darker, that is, the flame of the lamp had absorbed from the more brilliant solar beam light of the same color as it would itself emit. The explanation of the dark lines of the solar spectrum is obvious. The light from the body of the sun gives a continuous spectrum like that of an incandescent solid or liquid. Somewhere in its course this light passes through an atmosphere of gases which absorbs from the solar beam such light as these gases would emit if they were self-luminous. Some of this absorption occurs in the earth's atmosphere, but most of it is known to occur in the atmosphere of the sun itself. By comparison of these dark lines with the spectra of various incandescent substances upon which we can experiment, the probable constitution of the sun is inferred.

CHAPTER VI.

ABSORPTION AND EMISSION.

314. **Effects of Radiant Energy.**—It has been stated that the solar spectrum, whether produced by means of a prism or by a grating, may, under certain conditions, give rise to heat, light, or chemical changes. It was formerly supposed that these were dne to three distinct agents emanating from the sun, giving rise to three spectra which were partially superposed. Numerous experiments show, however, that, at any place in the spectrum where light, heat, and chemical effects are produced, nothing which we can do will separate one of these effects from the others. Whatever diminishes the light at any part of the spectrum diminishes the heat and chemical effects also. Physicists are now agreed that all these phenomena are due to vibratory motions transmitted from the sun, which differ in length of wave, and which are separated by a prism, because waves differing in length are transmitted in the substance of the prism with different velocities. The effect produced at any place in the spectrum depends upon the nature of the surface upon which the radiations fall. On the photographic plate they produce chemical change, on the retina the sensation of light, on the thermopile the effect of heat. Only those waves of which the wave lengths lie between 3930 and 7600 tenth metres affect the optic nerve. Chemical changes and the effects of heat are produced by radiations of all wave lengths.

To produce any effect the radiations must be absorbed; that is, the energy of the ethereal vibrations must be imparted to the substance on which they fall, and cease to exist as radiant energy. The most common effect of such absorption is to generate heat, and there are some surfaces upon which heat will be generated by the absorption of ethereal waves of any length. Langley, by means of the bolometer, has been able to measure the energy throughout the spectrum, and has shown the existence of lines like the Fraunhofer lines, in the invisible spectrum below the red. He has demonstrated the existence, in the lunar spectrum, of waves as long as 170,000 tenth metres, or more than twenty-two times as long as the longest that can excite human vision.

315. Intensity of Radiations.—The intensity of radiations can only be determined by their effects. If the radiations fall on a body by which they are completely absorbed and converted into heat, the amount of heat developed in unit time may be taken as the measure of the radiant energy. Let us suppose the radiations to emanate from a point equally in all directions, and represent the total intensity of the radiations by E. Let the point be at the centre of a hollow sphere, of which the radius is r, and represent by I the intensity of the radiations per unit area of the sphere. Then, since the surface of the sphere equals $4\pi r^2$, we have

$$E = 4\pi r^2 I,$$

and

$$I = \frac{E}{4\pi r^2}. \tag{119}$$

That is, the intensity of the radiation upon a given surface is in the inverse ratio of the square of its distance from the source.

If the surface is not normal to the rays, the radiant energy it receives is less, as will appear from Fig. 152. Let ab be a surface the normal to which makes with the ray the angle θ; then ab will receive the same quantity of radiant energy as $a'b'$, its projection on the plain normal to the ray. But $a'b'$ equals $ab \cos \theta$; and if I represent the intensity on $a'b'$, and I' the intensity on ab, we have

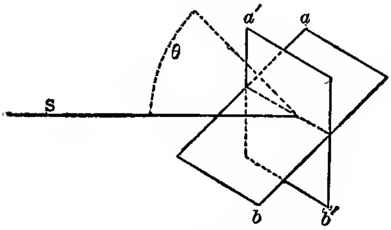

FIG. 152.

$$I' = I \cos \theta;$$

or, the intensity of the radiations falling on a given surface is proportional to the cosine of the angle made by the surface and the plane normal to the direction of the rays.

316. Photometry.—The object of *photometry* is to compare the luminous effects of radiations. It is not supposed that the radiations which fall on the retina are totally absorbed by the nerves that impart the sensation of light. The luminous effects, therefore, depend on the susceptibility of these nerves, and can only be compared, at least when different wave lengths are concerned, by means of the eye itself. The photometric comparison of two luminous sources is effected by so placing them that the illuminations produced by them respectively, upon two surfaces conveniently placed for observation, appear to the eye to be equal. If E and E' represent the intensities of the sources, I and I' the intensities of the illuminations produced by them on surfaces at distances r and r', the ratio between these intensities, as was seen in the last section, is

$$\frac{I}{I'} = \frac{\frac{E}{r^2}}{\frac{E'}{r'^2}} = \frac{Er'^2}{E'r^2};$$

and when I and I' are equal,

$$Er'^2 = E'r^2,$$

or

$$\frac{E}{E'} = \frac{r^2}{r'^2}. \qquad (120)$$

That is, when two luminous sources are so placed as to produce equal illuminations on a surface, their intensities are as the squares of their distances from the illuminated surfaces.

Rumford's photometer consists of a screen in front of which is an upright rod. The luminous sources are so placed that the rod casts two shadows near together upon the screen, and are adjusted at such distances that these shadows are apparently equal in intensity.

In *Foucault's photometer* the screen is of ground glass, and in place of the rod a vertical partition is placed in front of and perpendicular to the middle of the screen. The luminous sources are so placed that one illuminates the screen on one side of the partition, and the other on the other. The partition may be moved to or from the screen until the two illuminated portions just meet without overlapping.

In *Bunsen's photometer* the sources to be compared are placed on the opposite sides of a paper screen, a portion of which has been rendered translucent by oil or paraffine. When this screen is illuminated upon one side only, the translucent portion appears darker on that side, and lighter on the other side, than the opaque portion. When placed between two luminous sources, both sides of it may, by moving it toward one or the other, be made to appear alike, and the translucent portion almost invisible. The light transmitted through this portion in one direction then equals that transmitted in the opposite direction; that is, the two surfaces are equally illuminated.

317. Transmission and Absorption of Radiations.—It is a familiar fact that colored glass transmits light of certain colors only, and the inference is easy that the other colors are absorbed by the glass. It is only necessary to form a spectrum, and place the colored glass in the path of the light either before or after the separation of the colors, to show which colors are transmitted, and which absorbed.

By the use of the thermopile or bolometer, both of which are sensitive to radiations of all periods of vibration, it is found that some bodies are apparently perfectly transparent to light, and opaque to the obscure radiations. Clear, white glass is opaque to a large portion of the obscure rays of long wave length. Water and solution of alum are still more opaque to these rays, and pure ice transmits almost none of the radiations of which the wave lengths are longer than those of the visible red. Rock salt transmits well both the luminous and the non-luminous radiations.

On the other hand, some substances apparently opaque are transparent to radiations of long wave length. A plate of glass or rock salt rendered opaque to light by smoking it over a lamp is still as transparent as before to the radiations of longer wave length. Selenium is opaque to light, but transparent to the radiations of longer wave length. This fact explains the change of its electrical resistance by light, but not by non-luminous rays. Carbon disulphide, like rock salt, transmits nearly equally the luminous and non-luminous rays; but if iodine be dissolved in it, it will at first cut off the luminous rays of shorter wave length, and as the solution becomes more and more concentrated the absorption extends down the spectrum to the red, and finally all light is extinguished, and the solution to the eye becomes opaque. The radiations of which the wave lengths are longer than those of the red still pass freely. Black vulcanite seems perfectly opaque, yet it also transmits radiations of long wave length.

If the radiations of the electric lamp be concentrated by means of a lens, and a sheet of black vulcanite placed between the lamp and the lens, bodies may be still heated in the focus.

318. **Colors of Bodies.**—Bodies become visible by the light which comes from them to the eye, and bodies which are not self-luminous must become visible by sending to the eye some portion of the light that falls on them. Of the light which falls on a body, part is reflected from the surface; the remainder which enters the body is, in general, partly absorbed, and the unabsorbed portion either goes on through the body, or is turned back by reflection at a greater or less depth within the body, and mingles with the light reflected from the surface.

In general the surface reflection is small in amount, and the different colors are reflected almost in the proportion in which they exist in the incident light. Much the larger portion of the light by which a body becomes visible is turned back after penetrating a short distance beneath the surface, and contains those colors which the substance does not absorb. This determines the color of the object. In a few instances there seems to be a selective reflection from the surface. For example, the light reflected from gold-leaf is yellow, while that which it transmits is green.

319. **Absorption by Gases.**—If a pure spectrum be formed from the white light of the electric lamp, and sodium vapor, obtained by heating a bit of sodium or a bead of common salt in the Bunsen flame, be placed in the path of the beam, two narrow, sharply defined dark lines will be seen to cross the spectrum in the exact position that would be occupied by the yellow lines constituting the spectrum of sodium vapor. Gases in general have an effect similar to that of the vapor of sodium; that is, they absorb from the light which passes through them distinct radiations corresponding to definite wave lengths, which are always the same as those which would be emitted by the

gas were it rendered incandescent. It has been seen already (§ 313) that the Fraunhofer lines of the solar spectrum are thus accounted for.

320. Emission of Radiations.—Not only do incandescent bodies emit radiations, but all bodies at whatever temperature they may be. A warm body continues to grow cool until it arrives at the temperature of surrounding bodies, and then if it be moved to a place of lower temperature, it cools still further. To this process we can ascribe no limit, and it is necessary to admit that the body will radiate heat, and so grow cooler, whatever its own temperature, if only it be warmer than surrounding bodies. But it cannot be supposed that a body ceases to radiate heat when it comes to the temperature of surrounding bodies, and begins again when the temperature of these is lowered. It is necessary, therefore, to assume that all bodies at whatever temperature are radiating heat, and that, when any one of them arrives at a stationary temperature, it is, if no change take place within it involving the generation or consumption of heat, receiving heat as rapidly as it parts with it. This is called the principle of *movable equilibrium of temperature*. We know that if a number of bodies, none of which are generating or consuming heat otherwise than in change of temperature, be placed in an inclosure the walls of which are maintained at a constant temperature, these bodies will in time all come to the temperature of the inclosure. It can be shown that, for this to be true, the ratio of the emissive to the absorbing power must be the same for all bodies, not only for the sum total of all radiations, but for radiations of each wave length. For example, a body which does not absorb radiations of long wave length cannot emit them, otherwise, if placed in an inclosure where it could only receive such radiations, it would become colder than other bodies in the same inclosure. This is only a general statement of the fact which has been already stated for gases, that bodies absorb radiations of exactly the same kind as those which they emit.

Since radiant energy is energy of vibratory motion, it may be supposed to have its origin in the vibrations of the molecules of the radiating body. In § 156 it was shown that the various phenomena of gases are best explained by assuming a constant motion of their molecules. If these molecules should have definite periods of vibration, remaining constant for the same gas through wide ranges of pressure and temperature, this would fully explain the peculiarities of the spectra of gases.

In § 261 it was seen that a vibrating body may communicate its vibrations to another body which can vibrate in the same period, and will lose just as much of its own energy of vibration as it imparts to the other body. Moreover, a body which has a definite period of vibration is undisturbed by bodies vibrating in a period different from its own. This explains fully the selective absorption of a gas. For, if a beam of white light pass through a gas, there are, among the vibrations constituting such a beam, some which correspond in period to those of the molecules of the gas, and, unless the energy of vibration of these molecules is already too great, it will be increased at the expense of the vibrations of the same period in the beam of light. Hence, at the parts in the spectrum where light of those vibration periods would fall, the light will be enfeebled, and those parts will appear, by contrast, as dark lines.

In solids and liquids, the molecules are so constrained in their movements that they do not vibrate in definite periods. Vibrations of all periods may exist; but if in a given case there were a tendency to one period of vibration more than to another, it is evident that the body would transfer to or receive from another, that is, it would emit or absorb, vibrations of that period more than of any other. Furthermore, a good radiator is a body so constituted as to impart to the medium around it the vibratory motion of its own molecules. But the same peculiarity of structure which fits it for communicating its own motion to the medium when its own motion is the greater, fits it also for receiving motion from the medium when its own motion is the less. Theory, therefore, leads us to the conclusion which experiment has established, that at a given temperature emissive and absorbing powers have the same ratio for all bodies.

321. Loss of Heat in Relation to Temperature.—The loss of heat by a body is the more rapid the greater the difference of temperature between it and surrounding bodies. For a small difference of temperature the loss of heat is nearly proportional to this difference. This law is known as *Newton's law of cooling*. For a large difference of temperature the loss of heat increases more rapidly than the difference of temperature, and depends not merely upon this difference, but upon the absolute temperature of the surrounding bodies. An extended series of experiments by Dulong and Petit led to a formula expressing the quantity of heat lost by a body in an inclosure during unit time. It is

$$Q = m(1.0077)^{\theta}(1.0077^{t} - 1),$$

where θ represents the temperature of the inclosure, t the difference of temperature between the inclosure and the radiating body, both measured in Centigrade degrees, and m a constant depending on the substance, and the nature of its surface.

322. Kind of Radiation as Dependent upon Temperature.—When a body is heated we may feel the radiations from its surface long before those radiations render the body visible. If we continue to raise the temperature, after a time the body becomes red hot; as the temperature rises still further it becomes yellow, and finally attains a white heat. Even this rough observation indicates that the radiations of great wave length are the principal radiations at the lower temperature, and that to these are added shorter and shorter wave lengths as the temperature rises. Draper showed that the spectrum of a red-hot body exhibits no rays of shorter wave length than the red, but that as the temperature rises the spectrum is extended in the direction of the violet, the additions occurring in the order of the wave lengths. At the same time the colors previously existing increase in brightness, indicating an increase in energy of the vibrations of longer wave length as those of shorter wave length become visible. Experiments by Nichols on the radiations from glowing platinum show that vibrations of shorter wave length are not altogether absent from the radiations of a body of comparatively low temperature, and he was led to believe that all wave lengths are present in the radiations from even the coldest bodies, but are too feeble to be detected.

With gases, as has been seen, the radiations are apparently confined to a few definite wave lengths, but careful observations of the spectra of gases show that the lines are not defined with absolute sharpness, but fade away, although very rapidly, into the dark background. In many cases the existence of radiations may be traced throughout the spectrum, and it is a question whether the spectra of gases are not after all continuous, only showing strongly marked and sharply defined maxima where the lines occur. In general, increase of temperature does not alter the spectra of gases except to increase their intensity, but there are some cases in which additional lines appear as the temperature rises, and a few cases in which the spectrum undergoes a complete change at a certain temperature. This occurs with those compound gases which suffer dissociation at a certain temperature, and at higher temperatures give the spectra of their elements. When it occurs with gases supposed to be elements it suggests the question whether they are not really compounds, the molecules of which at the high temperature are divided, giving new molecules of which the rates of vibration are entirely different from those of the original body.

323. Fluorescence and Phosphorescence.—A few substances, such as sulphate of quinine, uranium glass, and thallene, have the property, when illuminated by rays of short wave length, even by the invisible rays beyond the violet, of emitting light of longer wave length. Such substances are *fluorescent*. The light emitted by them, and the conditions favorable to their luminosity, have been studied by Stokes. It appears that the light emitted is of the same character, covering a considerable region of the spectrum, no matter what may be the incident light, provided this be such as to produce the effect at all. The light emitted is always of longer wave length than that which causes the luminosity.

There is another class of substances which, after being exposed to light, will glow for some time in the dark. These are *phosphorescent*. They must be carefully distinguished from such bodies as phosphorus and decaying wood, which glow in consequence of chemical action. Some phosphorescent substances, especially the calcium sulphides, glow for several hours after exposure.

324. Anomalous Dispersion.—As has been already stated, there is a class of bodies which show a selective absorption at their surfaces. The light reflected from such bodies is complementary to the light which they can transmit. Kundt, following up isolated observations of other physicists, has shown that all such bodies give rise to an *anomalous dispersion;* that is, the order of the colors in the spectrum formed by a prism of one of these substances is not the same as their order in the diffraction spectrum or in the spectrum formed by prisms of substances which do not show selective absorption at their surfaces. Solid fuchsin, when viewed by reflected light, appears green. In solution, when viewed by transmitted light, it appears red. Christiansen allowed light to pass through a prism formed of two glass plates making a small angle with each other, and containing a solution of fuchsin in alcohol. He found that the green was almost totally wanting in the spectrum, while the order of the other colors was different from that in the normal spectrum. In the spectrum of fuchsin the colors in order, beginning with the one most deviated, were violet, red, orange, and yellow. Other substances give rise to anomalous dispersion in which the order of the colors is different.

In order to account for these phenomena, the ordinary theory of light is extended by the assumption that the ether and molecules of a body materially interact upon one another, so that the vibrations in a light-wave are modified by the vibrations of the molecules of a transparent body through which light is passing. This hypothesis, in the hands of Helmholtz and Ketteler, has been sufficient to account for most of the phenomena of light.

CHAPTER VII.

DOUBLE REFRACTION AND POLARIZATION.

325. Double Refraction in Iceland Spar.—If refraction take place in a medium which is not isotropic, as has been assumed in the previous discussion of refraction, but eolotropic, a new class of phenomena arises. Iceland spar is an eolotropic medium by the use of which the phenomena referred to are strikingly exhibited. Crystals of Iceland spar are rhombohedral in form, and a crystal may be a perfect rhombohedron with six equal plane faces, each of which is a rhombus.

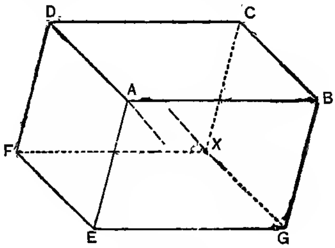

FIG. 153.

Fig. 153 represents such a crystal. At A and X are two solid angles formed by the obtuse angles of three plane faces. The line through A making equal angles with the three edges AB, AE, AD, or any line parallel to it, is an *optic axis* of the crystal.

Any plane normal to a surface of the crystal and parallel to the optic axis is called a *principal plane*. If such a crystal be laid upon a printed page, the lines of print will, in general, appear double. If a dot be made on a blank paper, and the crystal placed upon it, two images of the dot are seen. If the crystal be revolved about an axis perpendicular to the paper, one of the images remains stationary, and the other revolves around it. The images lie in a plane perpendicular to the paper, and parallel to the line joining the two obtuse angles of the face by which the light enters or emerges. The entering and emerging light is supposed in this case to be normal to the surfaces of the crystal. If the crystal be turned with its faces oblique to the light, the line joining the images will, in certain cases, not lie parallel to the line joining the obtuse angles of the faces. If the distances of the two images from the observer be carefully noticed it will be seen that the stationary one appears nearer than the other. If the obtuse angles A and X be cut away, and the new surfaces thus formed at right angles to the optic axis be polished, images seen perpendicularly through these faces do not appear double. By cutting the crystals into prisms in various ways its indices of refraction may be measured. It is found that, of the two beams into which light is, in general, divided in the crystal, one obeys the ordinary laws of refraction, and has a refractive index 1.658. It is called the *ordinary ray*. The other has no constant refractive index, does not in general lie in the normal plane containing the incident ray, and refraction may occur when the incidence is normal. It is the *extraordinary ray*. The ratio between the sines of the angles of incidence and refraction varies, for the Fraunhofer line D, from 1.658, the *ordinary index*, to 1.486. This minimum value is called the *extraordinary index*.

326. Explanation of Double Refraction.—In § 279 it was seen that the index of refraction of a substance is the reciprocal of the ratio of the velocity of light in the substance to its velocity in a vacuum. It is plain, then, that the velocity of light for the ordinary ray of the last section is the same for all directions, and, if light emanate from a point within the crystal, the light, following the ordinary laws of refraction, must proceed in spherical waves about that point as a centre, as in any single refracting medium. The phenomena presented by the extraordinary light in Iceland spar are fully explained by assuming that the velocities in different directions in the crystal are such as to give a wave front in the form of a flattened spheroid, of which the polar diameter, parallel to the optic axis, is equal to the diameter of the ordinary spherical wave, and the equatorial

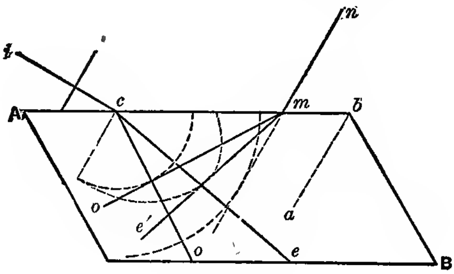

FIG. 154.

diameter is to its polar diameter as 1.658 is to 1.486. From these two wave surfaces the path of the light may easily be determined by construction by methods already explained in § 279, and exemplified in Fig. 154, in which *ic* represents the direction of the incident light, and *co* and *ce* the ordinary and extraordinary rays respectively.

327. Polarization of the Doubly Refracted Light.—If a second crystal be placed in front of the first in any of the experiments described in the last section, there will be seen in general four images instead of two; but if the second crystal be turned, the images change in brightness, and for four positions of the second crystal, when its principal plane is parallel or at right angles to the principal plane of the first, two of the images are invisible, and the other two are at a maximum brightness. If one of the beams of light produced by the first crystal be intercepted by a screen, and the other allowed to pass alone through the second crystal, the phenomena presented are easily followed. If the principal planes of the two crystals coincide, only one image is seen. If the second crystal be now rotated about the beam of light as an axis, a second image at once appears, at first very faint, but increasing in brightness. The original image at the same time diminishes in brightness, and the two are equally bright when the angle between the principal planes is 45°. If the angle be 90° the first image disappears, and the second is at its maximum brilliancy. As the rotation is continued the first image reappears, while the second grows dim and disappears when the angle between the principal planes is 180°. These changes show that the light which emerges from the first crystal of spar is not ordinary light. Another experiment shows this in a still more striking manner. Let the extraordinary ray be cut off by a screen, and the ordinary ray be received on a plane unsilvered glass at an angle of incidence of 57°. When the plane of incidence coincides with the principal plane of the spar, the light is reflected like ordinary light. If the mirror be now turned about the incident ray as an axis, that is, so turned that, while the angle of incidence remains unchanged, the plane of incidence makes successively all possible angles with the principal plane of the crystal, the reflected light gradually diminishes in brightness, and when the angle between the plane of incidence and the principal plane of the crystal is 90° it fails altogether. If the rotation be continued it gradually returns to its original brightness, which it attains when the angle between the same planes is 180°, and then diminishes until it fails when the angle is 270°. The extraordinary ray presents the same phenomena except that the reflected light is brightest when the angle between the planes is

90° and 270°, and fails when that angle is 0° and 180°. Beams of light after double refraction present different properties on different sides, and are said to be ***polarized.*** The explanation must, of course, be found in the character of the vibratory motion.

In the polarized beam it is plain that the vibrations must be transverse; for if the light were the result of longitudinal vibrations, or even of vibrations having a longitudinal component, it could not be completely extinguished for certain *azimuths* of the second crystal or of the glass reflector. The difference between ordinary and polarized light is explained if we assume that in both the vibrations of the ether particles take place at right angles to the line of propagation of the wave, and that in ordinary light they occur successively in all azimuths about that line, and may be performed in ellipses or circles as well as in straight lines, while in polarized light they occur in one plane. In the ordinary ray in Iceland spar the vibrations are in a plane at right angles to the optic axis. In the extraordinary ray they are in the plane containing the optic axis and the ray. The equation $v = \sqrt{\frac{E}{D}}$ holds for transverse vibrations, if by E be understood the modulus of rigidity of the medium. If we assume that the modulus of rigidity at right angles to the optic axis is a minimum, and along the optic axis a maximum, and varies between these two directions according to a simple law, all the phenomena of double refraction and polarization in the crystal are accounted for. If a crystal be cut so as to present faces parallel to the optic axis, and if light enter along a normal to one of these faces, the vibrations, which previous to entering the crystal were in all azimuths, are resolved in it in two directions, that of greatest and that of least elasticity, or parallel to and at right angles to the optic axis. The wave made up of vibrations parallel to the optic axis is propagated with the greater velocity. In this case the two wave fronts continue in parallel planes, and upon emergence constitute apparently one beam of light. If the incidence be oblique and in a plane at right angles to the principal plane, the two component vibrations are still parallel to and at right angles to the optic axis, but refraction occurs which is greater for the ray of which the vibrations are in the direction of least elasticity. If the incidence be oblique and in the principal plane, it is evident that there may be a component vibration at right angles to the optic axis, but the other component, since it must be at right angles to the ray, cannot be parallel to the optic axis, and therefore cannot be in the direction of greatest elasticity in the crystal. The second component is, however, in the direction of greatest elasticity in the plane of vibration, which direction is at right angles to the first component. In general, if a ray of light pass in any direction within the crystal, the line drawn at right angles to that direction and to the optic axis, that is, at right angles to the plane determined by the ray and the optic axis, is in the direction of least elasticity. One of the component vibrations is in that direction. A line drawn at right angles to the ray and in the plane formed by it and the optic axis is in the direction of the greatest elasticity to which any vibration giving rise to that ray of light can correspond. In that direction is the second component vibration. The two component vibrations are therefore always at right angles. One of the components is always at right angles to the optic axis, and hence in the direction of least elasticity. The light resulting from this component always travels with the same velocity whatever its direction, and hence suffers refraction on entering the crystal or emerging from it, according to the ordinary law for single refraction. The other component, being in the plane containing the ray and the optic axis and at right angles to the ray, may make all angles with the optic axis from 0° when it is in the direction of maximum elasticity and is propagated with the greatest velocity, to 90° when it is in a direction in which the elasticity is the same as that for the other component, and the entire beam is propagated as ordinary light. Light for which vibrations occur in all azimuths will, on entering the crystal, give rise to equal components, but light already polarized will give rise to components the intensities of which are determined by the law for the resolutions of motions. When its own direction of vibration coincides with that of either of the components, the other component will be zero, and only when its vibrations make an angle of 45° with the components can these components be equal. The varying intensities of the two beams into which a polarized beam is divided by a second crystal are thus explained.

328. Polarization by Reflection.—Light reflected from a transparent medium is found in general to be partially polarized, and for a certain angle of incidence the polarization is perfect. This angle is that for which the reflected and refracted rays are at right angles. In Fig. 155 let xy represent the surface of a transparent medium, ab the incident, bc the reflected, and bd the refracted ray. If the angle $cbd = 90°$, we have $r + i = 90°$ also; and since $\mu = \frac{\sin i}{\sin r}$, we have $\mu = \frac{\sin i}{\cos i} = \tan i$. Hence the angle of complete polarization is given by the equation $\tan i = \mu$. The fact embodied in this equation was discovered by Brewster, and is known as *Brewster's law.* The angle of complete polarization is called the *polarizing angle.* The plane of incidence is the *plane of polarization.* The vibrations of polarized light are at right angles to the plane of polarization. In the transmitted ray is an equal amount of polarized light the vibrations of which are in the plane of incidence.

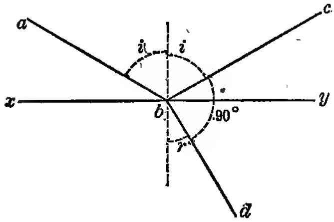

FIG. 155.

If a beam of ordinary light traverse a transparent medium, in which are suspended minute solid particles, the light which is reflected from them is found to be partially polarized. The maximum polarization is found in the light reflected at right angles to the beam. The plane of polarization of the polarized beam is the plane of the original beam and the beam which reaches the eye of the observer.

329. Polariscopes.—In experimenting with polarized light we need a *polarizer* to produce the polarized beam, and an *analyzer* to show the effects of the polarization. A piece of plane glass, reflecting light at the polarizing angle, is a simple polarizer. Double refracting crystals, if means be employed

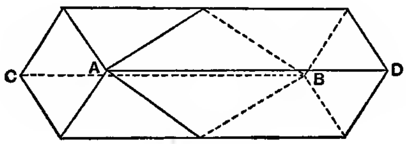

Fig. 156.

to suppress one of the beams into which the light is divided, are excellent polarizers. Tourmaline is a double refracting crystal which has the property of being more transparent to the extraordinary than to the ordinary ray. By grinding plates of tourmaline to the proper thickness, the ordinary ray is completely absorbed, while the extraordinary ray is transmitted. The best method of obtaining a polarized beam is by the use of a crystal of Iceland spar in which, by an ingenious device, the ordinary ray is suppressed, and the extraordinary transmitted. Fig. 156 shows how this is accomplished. *AB* is a crystal of considerable length. It is divided along the plane *AB* making an angle of 22° with the edge *AD* and perpendicular to a principal plane of the face *AC*. The faces of the cut are polished and the two halves cemented together again by Canada balsam.

in the same position as at first. In Fig. 157, which is a section through *ACBD* of Fig. 156, *ab* represents the direction of the light which is incident upon the face *AC*. It is separated into the two rays *o* and *e*. Since the refractive index of the balsam is intermediate between the ordinary and extraordinary indices of the spar, and since the angle *DAB* is so chosen that

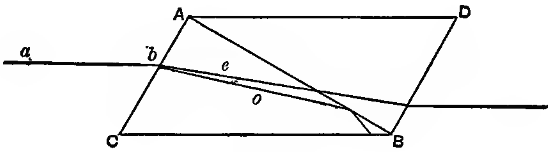

FIG. 157.

the ray *o* strikes the balsam at an angle of incidence greater than the critical angle, the ray *o* is totally reflected. The ray *e*, on the other hand, having a refractive index in the spar less than in the balsam, is not reflected, but continues through the crystal. A crystal of Iceland spar so treated is called a *Nicol's prism*, or often simply a *Nicol.*

The *Foucault prism* is similar to the Nicol, except that the two halves after polishing are not cemented together, but are mounted with a film of air between. The total reflection of *o* now occurs at a much less angle of incidence. The section *AB* is, therefore, much less oblique, and a shorter crystal serves for the construction of the prism. It will be observed that the section *AB* must be so made that the angle of incidence of *o* shall be greater, and of *e* less, than the corresponding critical angle. Since the two critical angles are nearly the same, but little variation in the angle of incidence of *o* and *e* is permissible, and the Foucault prism is, therefore, only useful for parallel rays.

A pair of Nicol's prisms, mounted with their axis coinciding, serve as a *polariscope*. The first Nicol transmits a single beam of polarized light the vibrations of which are in the principal plane. When the principal plane of the second Nicol coincides with that of the first this light is wholly transmitted through it. If the second Nicol or analyzer be turned about its axis, whenever its principal plane makes an angle with the direction of the vibrations, these are resolved into two components, one in and the other at right angles to the principal plane. The latter is reflected to one side and absorbed, and the former is transmitted. As the angle between the two principal planes increases, the transmitted component diminishes in intensity, until when this angle becomes 90° it disappears entirely. In this position the polarizer and analyzer are said to be *crossed*.

330. Effects of Plates of Doubly Refracting Crystals on Polarized Light.—If a plate cut from a doubly refracting surface so that its faces are parallel to the optic axis, or at least not at right angles to it, be placed between the crossed polarizer and analyzer, if the principal plane of the plate coincides with, or is at right angles to, the plane of vibration, no effect is perceived. But if the plate be rotated so that its principal plane makes an angle with the plane of vibration, the motion may be considered to be resolved into two components, one in, and the other at right angles to, the principal plane of the plate, and these two components on reaching the analyzer are again resolved each into two others, one in, and the other at right angles to, the principal plane of the analyzer. The vibrations in the principal plane of the analyzer are transmitted through it, and hence, in general, the introduction of the plate restores the light which the crossed polarizer and analyzer had extinguished. It is easy to see that the restored light will be most intense when the principal plane of the plate makes an angle of 45° with the plane of vibration of the polarized ray.

It is not to be understood that in the plate there are two separate beams of light, in one of which one set of particles is vibrating in one plane, and in the other another set in another plane. What really takes place is that each particle in the path of the light describes a path which is the resultant of the two components spoken of above. Let *ab*, Fig. 158, be a plate of Iceland spar, and *cd* the direction of its optic axis. Suppose the path of the light perpendicular to the plane of the paper, and *ef* to represent the direction of the disturbance produced by the entrance of a plane polarized wave. A motion in the direction of *ef* is compounded of two motions, one along the axis, and the other perpendicular to it. In the propagation of this motion to the next particle, the motion in the direction of the optic axis will begin a little sooner than that at right angles because of the greater elasticity in the former direction, and this difference becomes greater as the light is propagated into the plate. This is equivalent to a change in the relative phases of two vibrations at right angles, and this causes the path of a vibrating particle to change from the straight line to an ellipse. The result is, therefore, that, when the initial disturbance has any direction except in or at right angles to the principal plane of the plate, the motion of the vibrating particles within the plate becomes elliptical, the ellipses changing form as the distance from the front surface of the plate increases. It is entirely admissible, however, in the discussion of the problem to substitute for the actual motion its two components, as was done above.

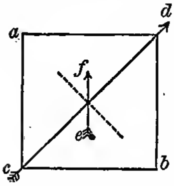

FIG. 158.

It remains to consider what is the effect of the retardation or change of phase of one of the components with respect to the other. It will be remembered that in the analyzer each ray from the plate is again resolved into two components, and that two of these components are in the principal plane of the analyzer and are transmitted. These two components will evidently differ in phase just as did the two motions from which they were derived, and since they are in the same plane their resultant is represented by their algebraic sum. If they differ in phase by half a period their algebraic sum will be zero, and no light will be transmitted by the analyzer. This will occur for a certain thickness of the interposed plate. If the light experimented upon be white, it may occur for some wave lengths and not for others. Hence, some of the constituents of white light may fail in the beam transmitted by the analyzer, and the image of the plate will then appear colored. A study of the resolution of the vibrations for this case shows that, of the two beams formed in the analyzer, one contains just that portion of the light that the other lacks; hence if the analyzer be turned through 90°, the image will change to the complementary color. In Fig. 159, let *ab* represent the plane of the vibrations in the polarized ray, and let *cd* and *ef* represent the two planes of vibration of the rays in the interposed plate. At the instant of entering the plate, the primary vibration and its two components will have the relation shown in the figure. The two components are then in the same phase. As the movement penetrates the plate, one component falls behind the other, and the relation of their phases changes, until, with a retardation of one wave length, the phases are again as in the figure. Suppose the thickness of the plate such that this retardation occurs for some constituent of white light. After leaving the plate the relative phases of the components remain unchanged and the constituent in question enters the analyzer as two vibrations at right angles and in the same phase. In Fig. 160, let *oe* and *od* represent the two components, and *xx* and *yy* the two planes of vibration in the analyzer. *oe* will give the components *om* and *on*, and *od* the components *om'* and *on'*. Since the components *om* and *om'* annul one another, the color to which they correspond is wanting in the light resulting from vibrations in

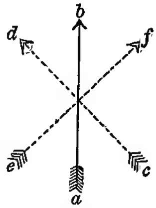

FIG. 159.

the plane xx, while since the components on and on' are added, this color is found in full intensity among the vibrations in the plane yy. For light of other wave lengths, the relative retardation is different, but for each vibration period, the component in the direction xx combined with that in the direction yy represents the total light for that period in the beam entering the analyzer; that is, the total effect of vibrations in the direction xx combined with that of vibrations in the direction yy must produce white light, and one effect must, therefore, be the complement of the other.

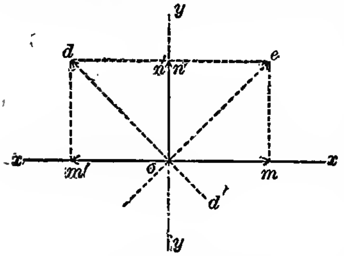

FIG. 160.

Let us suppose the plate thick enough to cause a retardation equal to a certain number of wave lengths, which we will assume to be ten, of the shortest waves of the visible spectrum. Since the longest waves of the visible spectrum are about twice the length of the shortest, they will suffer a retardation of five wave lengths. Other waves will suffer a retardation of nine, eight, seven, and six wave lengths. But, as was seen above, a retardation of one or more whole wave lengths of any kind of light causes extinction of that kind of light in the beam transmitted by the crossed analyzer. In the case considered the transmitted beam will lose six kinds of light distributed at about equal distances along the spectrum. The light remaining will consist of the different colors in about the same proportions as they exist in white light, and the beam will therefore be white but diminished in intensity. Hence, when a thick plate is interposed between the crossed polarizer and analyzer the restored light is white.

331. Elliptic and Circular Polarization.—In the last section, in discussing the effects of a thin plate, we considered the two components of the vibratory motion propagated from it. It was stated that the real motion of the vibrating particles was in general elliptical. Let us consider more fully the real motion. Let us suppose that the light is light of one wave length only, and that, as before, the principal plane of the plate makes an angle of 45° with the plane of vibration of the incident light. In Fig. 161 let *yy* represent the original plane of vibration, and *ab* and *cd* the planes of maximum and minimum elasticity in the plate. As already explained, the first disturbance as the light enters the plate is in the direction *yy*; but as the disturbance is propagated into the plate, each disturbed particle receives an impulse first of all in the direction *cd* of greatest elasticity, then in other directions between *cd* and *ab*, and finally in the direction *ab*. From this results an elliptical orbit with the major axis in the direction *yy*. To determine this orbit exactly it is only necessary to take account of the time that elapses between the impulse in the direction *cd* and the corresponding impulse in the direction *ab*. It is sufficient to consider any particle as actuated by two vibratory motions in the directions *cd* and *ab* at right angles, and differing in phase. In Fig. 161, one side of the rectangle represents the greatest displacement in the direction *cd*, and the other side the displacement occurring at the same instant in the direction *ba*. The point *r* will represent the actual position of the vibrating particle. Constructing now the successive displacements of the particles in the directions *cd* and *ba* and combining these, we have the elliptical path as shown. As the light penetrates farther and farther into the plate the relative phases of the two vibrations change continually, and the ellipse passes through

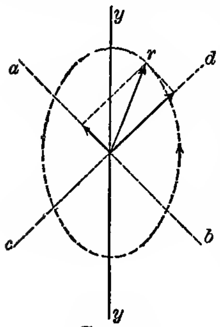

FIG. 161.

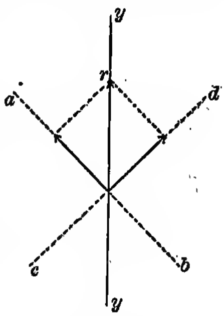

FIG. 162.

all its forms from the straight line *yy* to the straight line *xx* at right angles to it and back to the straight line *yy*. The direction of the path of the particle in the surface of the plate as the light emerges will be the direction of the path of all the particles in the polarized beam beyond the plate. If the component vibrations be in the same phase, that is, if they reach their elongations in the directions *ba* and *cd* (Fig. 162) at the same instant, the resultant vibration is in the line *yy* and the light is plane polarized exactly as it left the polarizer. This will occur when the retardation of light in the plane of *ba* with respect to that in the plane of *cd* is one, two, or more whole wave lengths. When the retardation is one half, three halves, or any odd number of half wave lengths, the phases of the two vibrations are as shown in Fig. 163, and the resultant is a plane polarized beam the vibrations of which are at right angles to those of the beam from the polarizer. A case of special interest is shown in Fig. 164, in which the difference of phase is one fourth a period, and the resultant vibration is a circle. A difference of three fourths will give a circle also, but with the rotation in the opposite direction.

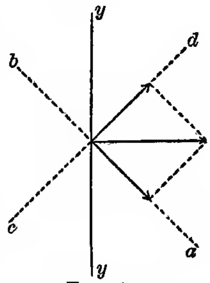

Fig. 163.

A plate of such thickness as to produce a retardation of one quarter of a wave length will give a circular vibration, and the beam issuing from the plate is then *circularly polarized*. Its peculiarity is that the two beams into which it is divided by a double refracting crystal are always of the same intensity, and no form of analyzer will distinguish it from ordinary light. *Quarter wave plates* are often made by splitting sheets of mica until the required thickness is obtained.

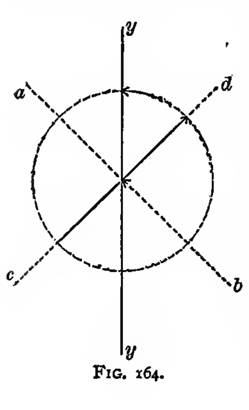

Fig. 164.

332. **Circular Polarization by Reflection.**—It has been seen that light reflected from a transparent medium at a certain angle is polarized, and that an equal amount of polarized light exists in the refracted beam. Light totally reflected in the interior of a medium is also polarized, and here, there being no refracted beam, the two components exist in the reflected light, but so related in phase that the light is elliptically polarized. Fresnel has devised an apparatus known as *Fresnel's rhomb,* by means of which circularly polarized light is obtained by two internal reflections of a beam of light previously polarized in a plane at an angle of 45° with the plane of incidence.

333. **Effect of Plates Cut Perpendicularly to the Axis from a Uniaxial Crystal.**—A crystal, such as Iceland spar, which has but one optic axis, is called a *uniaxial crystal.* Polarized light passing perpendicularly through a plate cut from such a crystal perpendicularly to its optic axis suffers no change. If, however, the plate between the crossed polarizer and analyzer be inclined to the direction of the beam, light passes through the analyzer. It is generally colored, the color changing with the obliquity of the plate. If a system of lenses be used to convert the polarized beam into a conical pencil and the plate be placed in this perpendicular to its axis, the central ray of the pencil will be unchanged, but the oblique rays will be resolved except in and at right angles to the plane of vibration, and there will appear beyond the analyzer a system of colored rings surrounding a dark centre, and intersected by a black cross. If the analyzer be turned through 90°, a figure complementary to the first in all its shades and tints is obtained; the black cross and centre become white, and the rings change to complementary colors.

334. **Biaxial Crystals.**—Most crystals have two optic axes or lines of no double refraction, instead of one. They are *biaxial crystals.* Their optic axes may be inclined to each other at any angle from 0° to 180°. The wave surfaces within

these crystals are no longer the sphere and the ellipsoid, but surfaces of the fourth order with two nappes tangent to each other at four points where they are pierced by the optic axes. Neither of the two rays in such a crystal follows the law of ordinary refraction. The outer wave surface around one of the points of tangency has a depression something like that of an apple around the stem. By reference to the method already employed for constructing a wave front, it will be seen that there may be such a position for the incident wave that, when the elementary wave surfaces are constructed, the resultant wave will be a tangent to them in the circle around one of these depressions where it is pierced by the optic axis. Now since the direction of a ray of light is from the centre of an elementary wave surface to the point of tangency of that surface and the resultant wave, we shall have in this case an infinite number of rays forming a cone, of which the base is the circle of tangency. In other words, one ray entering the plate in a proper direction may be resolved into an infinite number of rays forming a cone, which will become a hollow cylinder of light on emerging from the crystal. This phenomenon is called *conical refraction.* It was predicted by Hamilton from a mathematical analysis of the wave propagation in such crystals.

If a plate be cut from a biaxial crystal perpendicular to the line bisecting the angle formed by the optic axes, and placed between the polarizer and analyzer in a conical pencil of light, there will be seen a series of colored curves called lemniscates, resembling somewhat a figure 8. The existence of this phenomenon was also predicted and the forms of the curves investigated by mathematical analysis before they were seen.

335. Double Refraction by Isotropic Substances when Strained.—A piece of glass between the crossed polarizer and analyzer, if subjected to forces tending to distort it, will restore the light beyond the analyzer and in some cases produce chromatic effects. Unequal heating produces this result, and a long tube made to vibrate longitudinally shows it when the light crosses it near the node. Pieces cut from plates of unannealed glass exhibit double refraction when examined by polarized light. Indeed, the absence of double refraction is a test of perfect annealing.

336. Effects of Plates of Quartz.—A quartz crystal is uniaxial, and gives an ordinary and an extraordinary ray, but is unlike Iceland spar in that the extraordinary wave front in it is a prolate spheroid and lies wholly within the spherical ordinary wave, not touching it even where it is pierced by the optic axis. The effects due to plates of quartz in polarized light differ very greatly from those due to Iceland spar or selenite. If a plate of quartz cut perpendicular to the axis be placed in a beam of parallel, homogeneous, plane polarized light at right angles to its path, the light is, in general, restored beyond the analyzer, and is unchanged by the rotation of the quartz through any azimuth. If the analyzer be rotated through a certain angle, depending on the thickness of the quartz plate, the light is extinguished. It is evident that the plane of polarization has simply been rotated through a certain angle. Light of a different wave length would have been rotated through a different angle. A beam of white polarized light, therefore, has the planes of polarization of its constituents rotated through different angles, and the effect of rotating the analyzer is to quench one after another of the colors as the plane of polarization for each is reached. The result is a colored beam which changes its tint continuously as the analyzer rotates.

The best explanation of these phenomena was given by Fresnel. It is found that neither of the two beams from a quartz crystal is plane polarized. The polarization is in general elliptical, but becomes circular for waves perpendicular to the axis of the crystal, the motion in one ray being right-handed and in the other left-handed. Each particle of ether in the path of the light within the crystal is actuated at the same time by two circular motions in opposite directions. Its real motion is in the diameter which bisects the chord joining any two simultaneous hypothetical positions of the particle in the two circles. In Fig. 165 let P and Q represent these two simultaneous positions. It is plain that the two components in the direction AB have the same value and are added, while those at right angles to AB are equal and opposite and annul each other.

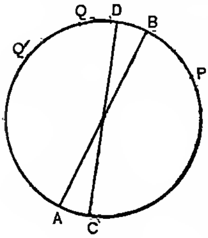

Fig. 165.

So long as the two components retain the same relation as that assumed, the real motion of the particle is in the line AB. But in the quartz plate one of the motions is propagated more rapidly than the other, and another particle farther on in the path of the light may reach the point P in one of its circular vibrations at the same time that it reaches Q' in the other. This will give CD as its real path, and the plane of its vibration has been rotated through the angle BOD. When the light finally emerges from the plate its plane of vibration will have been rotated through an angle which is proportional to the thickness of the plate and depends upon the wave length of the light employed. A plate of quartz one millimetre in thickness rotates the plane of polarization of red light corresponding to Fraunhofer's line B, 15° 18′, of blue light corresponding to the line G, 42° 12′. Some specimens of quartz rotate the plane of polarization in one direction, and some in the opposite. Rotation which is related to the direction of the light as the directions of rotation and propulsion in a right-handed screw is said to be *right-handed*, and that in the opposite direction is *left-handed*.

337. **Artificial Quartzes.**—Reusch has reproduced all the effects of quartz plates by superposing thin films of mica, each film being turned so that its principal plane makes an angle of 45° or 60°, always in the same direction, with that of

the film below. If a plane polarized wave enter such a combination, an analysis of the resolution of the vibration as it passes from film to film will show that the result is equivalent to that of two contrary circular vibrations, one of which is propagated less rapidly than the other. This helps to establish Fresnel's theory of the rotational effects of quartz.

338. **Rotation of the Plane of Polarization by Liquids.**—Many liquids rotate the plane of polarization, but to a less amount than quartz. A solution of sugar produces a rotation varying with the strength of the solution, and instruments called *saccharimeters* are made for determining the strength of sugar solutions from their effect in rotating the plane of polarization. In these instruments the effect is often measured by interposing a wedge-shaped piece of quartz, and moving it until a thickness is found which exactly compensates the rotation produced by the solution.

339. **Electromagnetic Rotation.**—Faraday discovered that when polarized light passes through certain substances in a magnetic field, the plane of polarization is rotated through a certain angle. The experiment succeeds best with a very dense glass consisting of borate of lead, so placed that the light may traverse it along the lines of magnetic force, in the field produced by a powerful electromagnet. The amount of rotation is proportional to the difference of magnetic potential between the two ends of the glass. The direction of rotation, as was shown by Verdet, is generally right-handed in diamagnetic media, and left-handed in paramagnetic media. It also depends upon the direction of the lines of force, and is therefore reversed by reversing the current in the electromagnet. It follows, also, that if the light, after traversing the glass with the lines of force, be reflected back through the glass against the lines of force, the rotation will be doubled. It is important to note that this is the reverse of the effect produced by quartz,

solutions of sugar, etc., which rotate the plane of polarization in consequence of their own molecular state. When light of which the plane of polarization has been rotated by passage through such substances is reflected back upon itself, the rotation produced during the first passage is exactly reversed during the return, and the returning light is found to be polarized in the same plane as at first.

In the magnetic field the effect is as though the medium which conveys the light were to rotate around an axis parallel to the lines of force, and to carry with it the plane of vibration. Evidently the plane of vibration would be turned through a certain angle during the passage of the light through the body, and would be turned still further in the same direction if the light were to return. An illustration may be drawn from the movement of a boat rowed across a current. If we row at right angles to the current, the boat is carried downward, and lands on the opposite shore below the point of starting. If then we row back, still at right angles to the current, the boat on reaching the shore from which it started is farther down the stream. On the other hand, in moving across a still lake, we might find ourselves compelled to take an oblique course on account of rocks or other permanent obstacles. If so, we should, on returning, be compelled to retrace our path, and would land at the point of starting.

When we remember that iron becomes magnetic by the effect of currents of electricity flowing in conductors around it, and that Ampère conceived that a permanent magnet consists of molecules surrounded by electric currents, all in the same direction, it is easy to imagine that the magnetic field is a region where the ether is actuated by vortical motions, all in the same direction, and in planes at right angles to the lines of magnetic force. Such a motion would account for the rotational effects of the magnetic field upon polarized light.

Not only glass but most liquids and gases exhibit rotational effects when placed in a powerful magnetic field; and Kerr has shown that when light is reflected from the polished pole of an electromagnet, its primitive plane of polarization is rotated when the current is passed, in one direction for a north pole, and in the opposite direction for a south pole.

340. **Maxwell's Electromagnetic Theory of Light.**—In Maxwell's treatment of electricity and magnetism, he assumed that electrical and magnetic actions take place through a universal medium. In order to determine whether this medium may not be identical with the luminiferous ether, he investigated its properties when a periodic electromagnetic disturbance is supposed to be set up in it, such as would result from a rapid reversal of electromotive force at a point, and compared them with the observed properties of the ether, on the assumption that light is an electromagnetic disturbance. He showed that such a disturbance would be propagated through the medium in a way similar to that in which vibrations are transmitted in an elastic solid. He showed further that if light were such a disturbance, its velocity in the ether should be equal to v, the ratio of the electrostatic to the electromagnetic system of units. Numerous measurements of the velocity of light and of this ratio show that they are very nearly equal.

He also showed that the indices of refraction of transparent media should be equal to the square roots of their specific inductive capacities. This relation may be deduced as follows: We may suppose electrical and luminous effects to be transmitted through the dielectric by means of the ether within it, and farther suppose electrical effects in the medium, and therefore its specific inductive capacity, to be proportional to displacements produced in the ether in it by electrical forces. Other things being equal, a displacement is inversely proportional to the elasticity of the medium. The velocity of propagation of a disturbance is directly proportional to the square root of the elasticity, if the density of the ether remain constant, and the index of refraction for light is inversely as the velocity of propagation. Hence the index of refraction is equal to the square root of the specific inductive capacity. To illustrate this let us suppose the specific inductive capacity of a dielectric to be 2. This means that a given electric force produces in the ether in that substance twice the displacement which it would produce in the ether in air. Hence the elasticity of the ether in that substance is one half as great as in air, the velocity of propagation of light in it will be to the velocity in air as $1 : \sqrt{2}$, and the index of refraction will be $\sqrt{2}$.

Measurements of indices of refraction and specific inductive capacities have shown that the relation which has been stated holds true in many cases. Hopkinson has shown, however, that it does not hold true for animal and vegetable oils.

The theory leads to the conclusion that the direction of propagation of the electrical disturbance and the accompanying magnetic disturbance at right angles to it is normal to the plane of these disturbances. By making the assumption, which is justified by Boltzmann's measurements upon sulphur, that an eolotropic medium has different specific inductive capacities in different directions, Maxwell shows also that the propagation of the electrical disturbance in a crystal will be similar to that of light. It has also been shown that the electrical disturbance will be reflected, refracted, and polarized at a surface separating two dielectrics.

Lastly, Maxwell concludes that, if his theory be true, bodies which are transparent to the vibrations of the ether should be dielectrics, while opaque bodies should be good conductors. In the former the electrical disturbance is propagated without loss of energy; in the latter the disturbance sets up electrical currents, which heat the body, and the disturbance is not propagated through the body. Observation shows that, in fact, solid dielectrics are transparent, and solid conductors are opaque, to radiations in the ether. Maxwell explains the fact that many electrolytes are transparent and yet are good conductors by supposing that the rapidly alternating electromotive forces which occur during the transmission of the electrical disturbance act for so short a time in one direction, that no complete separation of the molecules of the electrolyte is effected. No electrical current, therefore, is set up in the electrolyte, and electrical energy is not lost during the transmission of the disturbance.

29

TABLES.

TABLE I.

Units of Length.

Foot	=	30.48 cm.	log. 1.484015
Inch	=	2.54 cm.	log. 0.404830

Units of Mass.

Pound	=	453.59 grams.	log. 2.656664
Grain	=	0.0648 grams.	log. 8.811575

TABLE II.

Acceleration of Gravity.

$g = 980.6056 - 2.5028 \cos 2l - 0.000003h$, where l is the latitude of the station and h its height in centimetres above the sea level.

g at Washington	=	980.07	g at Paris	=	980.94
g at New York	=	980.26	g at Greenwich	=	981.17

TABLE III.

Units of Work.

Kilogram-metre	=	$100{,}000g$ ergs.
Foot-pound	=	$13{,}825g$ ergs.
	=	1.355×10^7 ergs, log 7.13200, when $g = 980$.

Units of Rate of Working.

Watt	=	10^7 ergs per second.
Horse-power	=	550 foot-pounds per second.
	=	746 Watts.

Unit of Heat.

Lesser calorie (gram-degree) $= 4.16 \times 10^7$ ergs.

TABLE IV.

Densities of Substances at 0°.

The densities of solids given in this table must be taken as only approximate. Specimens of the same substance differ among themselves to such an extent as to render it impossible to give more precise values.

Aluminium	2.6	Iron (wrought)	7.6 to 7.8
Brass	8.4	" (cast)	7.2 to 7.7
Copper	8.9	Lead	11.3
Gold	19.3	Mercury	13.596
Glass (crown)	2.5 to 2.7	Platinum	21.5
Hydrogen	0.0000895	Silver	10.5
Ice	0.918	Zinc	7.1

TABLE V.

Units of Pressure for $g = 981$.

	Grams per sq. cm.	Degrees per sq. cm.
Pound per square inch	70.31	6.9×10^4
1 inch of mercury at 0°	34.534	3.388×10^4
1 millimetre of mercury at 0°	1.3596	1333.8
1 atmosphere (760 mm.)	1033.3	1.0136×10^6
1 atmosphere (30 inches)	1036.	1.0163×10^6

TABLE VI.

Elasticity.

If p is the force in degrees per unit area tending to extend or compress a body, the linear elasticity is $\frac{dp}{dl}$, and the volume elasticity is $\frac{dp}{dv}$.

	$\frac{dp}{dl}$.	$\frac{dp}{dv}$.
Glass	6.03×10^{11}	4.15×10^{11}
Steel	2.14×10^{12}	1.84×10^{12}
Brass	1.07×10^{12}	....
Mercury	....	3.44×10^{10}
Water	....	2.02×10^{10}

TABLE VII.

Absolute Density of Water at $t°$ in Grams per Cubic Centimetre.

$t°$.	Density.	$t°$.	Density.	$t°$.	Density.
0	0.999884	7	0.999946	40	0.99236
1	0.999941	8	0.999899	50	0.98821
2	0.999982	9	0.999837	60	0.98339
3	1.000004	10	0.999760	70	0.97795
4	1.000013	15	0.999173	80	0.97195
5	1.000003	20	0.998272	90	0.96557
6	0.999983	30	0.995778	100	0.95866

TABLE VIII.

Density of Mercury at $t°$, Water at 4° being 1.

$t°$.	Density.	log.	$t°$.	Density.	log.
0	13.5953	1.13339	20	13.5461	1.13182
10	13.5707	1.13260	30	13.5217	1.13103

TABLE IX.

Coefficients of Linear Expansion.

	Temperature.	$a = \frac{dl}{dt}$.
Aluminium	16° to 100°	0.0000235
Brass	0 to 100	0.0000188
Copper	0 to 100	0.0000167
German silver	0 to 100	0.0000184
Glass	0 to 100	0.0000071
Iron	13 to 100	0.0000123
Lead	0 to 100	0.0000280
Platinum	0 to 100	0.0000089
Silver	0 to 100	0.0000194
Zinc	0 to 100	0.0000230

Coefficients of voluminal expansion, $\frac{dV}{dt} = 3\alpha$.

TABLE X.

SPECIFIC HEATS—WATER AT 0° = 1.

Solids and Liquids.

Aluminium	0.212	Mercury	0.033
Brass	0.086	Platinum	0.032
Copper	0.093	Silver	0.056
Iron	0.112	Water (0° to 100°)	1.005
Lead	0.031	Zinc	0.056

Gases and Vapors at Constant Pressure.

Air	0.237	Nitrogen	0.244
Hydrogen	3.410	Oxygen	0.217

$$\text{Ratio, } \frac{C_p}{C_v} = 1.404.$$

TABLE XI.

I. MELTING POINTS. II. BOILING POINTS. III. HEATS OF LIQUEFACTION. IV. HEATS OF VAPORIZATION. V. MAXIMUM PRESSURE OF VAPOR AT 0° IN MILLIMETRES OF MERCURY.

	I.	II.	III.	IV.	V.
Ammonia	..	− 33.7°	..	294 at 7.8°	3344
Carbon dioxide	− 65°	− 78.2	..	49.3 at 0°	27100
Chlorine	..	− 33.6	..	..	4560
Copper	1200	..	..	..	..
Lead	325	..	5.9°	..	..
Mercury	− 39	357	2.8	62	0.02
Nitrous oxide, N_2O	..	− 105	..	..	24320
Platinum	1780	..	27.2	..	..
Silver	1000	..	21.1	..	..
Water	0	100	80	537	4.6
Zinc	415	..	28.1	..	..

TABLE XII.

MAXIMUM PRESSURE OF VAPOR OF WATER AT VARIOUS TEMPERATURES IN (I.) DYNES PER SQUARE CENTIMETRE, (II.) MILLIMETRES OF MERCURY.

Temp.	I.	II.	Temp.	I.	II.
$-20°$	1236	..	60°	1.985×10^5	149.
$-10°$	2790	..	80	4.729×10^5	355.
0°	6133	4.6	100	10.14×10^5	760.
10	12220	9.2	120	19.88×10^5	1491.
20	23190	17.4	140	36.26×10^5	2718.
30	42050	31.5	160	62.10×10^5	4652.
40	73200	54.6	180	100.60×10^5	7546.
50	1.226×10^5	96.2	200	$156. \times 10^5$	11689.

TABLE XIII.

CRITICAL TEMPERATURES (T) AND PRESSURES IN ATMOSPHERES (P), AT THEIR CRITICAL TEMPERATURES, OF VARIOUS GASES.

	T.	P.		T.	P.
Hydrogen	$-174.$	99.	Carbon dioxide	30.9	77.
Nitrogen	$-124.$	42.	Sulphur dioxide	155.4	79.
Oxygen	$-105.$	49.			

TABLE XIV.

COEFFICIENTS OF CONDUCTIVITY FOR HEAT (K) IN C. G. S. UNITS, IN WHICH Q IS GIVEN IN LESSER CALORIES.

Brass	0.30	Mercury	0.015
Copper	1.11	Paraffin	0.00014
Glass	0.0005	Silver	1.09
Ice	0.0057	Vulcanized india-rubber	0.00009
Iron	0.16	Water	0.0015
Lead	0.08		

TABLE XV.

Energy Produced by Combination of 1 Gram of Certain Substances with Oxygen.

	Gram-degree of Heat.	Energy in ergs.
Carbon, forming CO	2141	8.98×10^{10}
" " CO_2	8000	3.36×10^{11}
Carbon monoxide, forming CO_2	2420	1.02×10^{11}
Copper, CuO	602	2.53×10^{10}
Hydrogen, H_2O	34000	1.43×10^{12}
Marsh gas, CO_2 and H_2O	13100	5.50×10^{11}
Zinc, ZnO	1301	5.46×10^{10}

TABLE XVI.

Atomic Weights and Combining Numbers.

	Atomic Weight.		Combining Number.
Aluminium	27.04		9.01
Copper	63.18	(cupric)	31.59
"	"	(cuprous)	63.18
Gold	196.2		65.4
Hydrogen	1.		1.
Iron	55.88	(ferric)	18.63
"	"	(ferrous)	27.94
Mercury	199.8	(mercuric)	99.9
"	"	(mercurous)	199.8
Nickel	58.6		29.3
Oxygen	15.96		7.98
Platinum	194.3		64.8
Silver	107.7		107.7
Zinc	64.88		32.44

TABLE XVII.

Molecular Weights and Densities of Gases.

Simple Gases.

	Atomic Weight.	Sp. gr., $H = 1$.	Mass in 1 litre.
Chlorine, Cl_2	70.75	35.37	3.167
Hydrogen, H_2	2.00	1.00	0.0895
Nitrogen, N_2	28.024	14.012	1.254
Oxygen, O_2	31.927	15.96	1.429

Compound Gases.

	Atomic Weight.	Sp. gr., $H = 1$.	Mass in litre.
Carbonic oxide, CO	27.937	14.97	1.251
Carbonic dioxide, CO_2	43.90	21.95	1.965
Hydrochloric acid, HCl	36.376	18.188	1.628
Vapor of water, H_2O	17.96	8.98	0.804
Atmospheric air			1.293

TABLE XVIII.

Electromotive Force of Voltaic Cells.

Daniell....... 1.1 volt. | Grove........ 1.88 volt. | Clark... 1.435 volt at 15°.

Electromotive force of Clark cell for any temperature t is

$$1.435[1 - 0.00077(t - 15)].$$

TABLE XIX.

Electro-chemical Equivalents.

Grams per second deposited by the electromagnetic unit current,

Hydrogen, 0.0001038.

To find the electro-chemical equivalents of other substances, multiply the electro-chemical equivalent of hydrogen by the combining number of the substance.

TABLE XX.

Electrical Resistance.

Absolute resistance R in C. G. S. units of a centimetre cube of the substance.

Temperature coefficient, α. $R_t = R_0(1 + \alpha t)$.

	R_0.	α.
Aluminium	2889	..
Copper	1611	0.00388
German silver	20763	0.00044
Gold	2041	0.00365
Iron	9638	..
Mercury	94340	0.00072
Platinum	8982	0.00376
Platinum silver, 2 Pt. 1 Ag.	24190	0.00031
Silver	1580	0.00377
Zinc	5581	0.00365

	R_0.
Carbon (Carré's electric light)..................	3.9×10^6
Glass at 200°....................................	2.23×10^{16}
Gutta-percha, at 24°.............................	3.46×10^{23}
" " " 0°...	6.87×10^{24}
Selenium, at 100°................................	5.9×10^{13}
Water, at 22°....................................	7.0×10^{10}
Zinc sulphate + 23 H_2O........................	1.83×10^{10}
Copper sulphate + 45 H_2O......................	1.91×10^{10}

TABLE XXI.

Indices of Refraction.

	Index.	Kind of Light.		Index.	Kind of Light.
Soft crown glass.....	1.5090	A	Canada balsam......	1.528	Red
	1.5180	E	Water...............	1.331	B
	1.5266	G		1.336	E
Dense flint glass.....	1.6157	B		1.344	H
	1.6289	E	Carbon disulphide....	1.614	A
	1.6453	G		1.646	E
Rock salt............	1.5366	A		1.684	G
	1.5490	E	Air at 0°, 760 mm....	1.00029	A
	1.5613	G		1.000296	E
Diamond...........	2.47	D		1.000300	H
Amber.............	1.532	D			

	Ordinary Index.	Kind of Light.	Extraordinary Index.
Iceland spar......	1.658	D	1.486
Quartz...........	1.544	D	1.553

TABLE XXII.

Wave Lengths of Light—Rowland's Determinations.

Fraunhofer's line A	(edge),	7593.975	tenth metres.
B	"	6867.382	
C	"	6562.965	
D_1	"	5896.080	
D_2	"	5890.125	
E	"	5270.429	
b	"	5183.735	
F	"	4861.428	
G	"	4307.961	

TABLE XXIII.

ROTATION OF PLANE OF POLARIZATION BY A QUARTZ PLATE, 1 MM. THICK, CUT PERPENDICULAR TO AXIS.

A	12°.668	E	27°.543
B	15°.746	F	32°.773
C	17°.318	G	42°.604
D_2	21°.727	H	51°.193

TABLE XXIV.

VELOCITIES OF LIGHT.

	Cm. per Sec.		Cm. per Sec.
Michelson, 1879	2.99910×10^{10}	Cornu, 1874	2.98500×10^{10}
Michelson, 1882	2.99853×10^{10}	Cornu, 1878	2.99990×10^{10}
Newcomb, 1882	2.99860×10^{10}	Young and Forbes, 1880	3.01382×10^{10}
Foucault, 1862	2.98000×10^{10}		

THE RATIO BETWEEN THE ELECTROSTATIC AND ELECTROMAGNETIC UNITS.

	Cm. per Sec.		Cm. per Sec.
Weber and Kohlrausch	3.1074×10^{10}	Exner	2.920×10^{10}
W. Thomson	2.825×10^{10}	Klemenčič	3.018×10^{10}
Maxwell	2.88×10^{10}	Himstedt	3.007×10^{10}
Ayrton and Perry	2.98×10^{10}	Colley	3.015×10^{10}
J. J. Thomson	2.963×10^{10}		

INDEX.